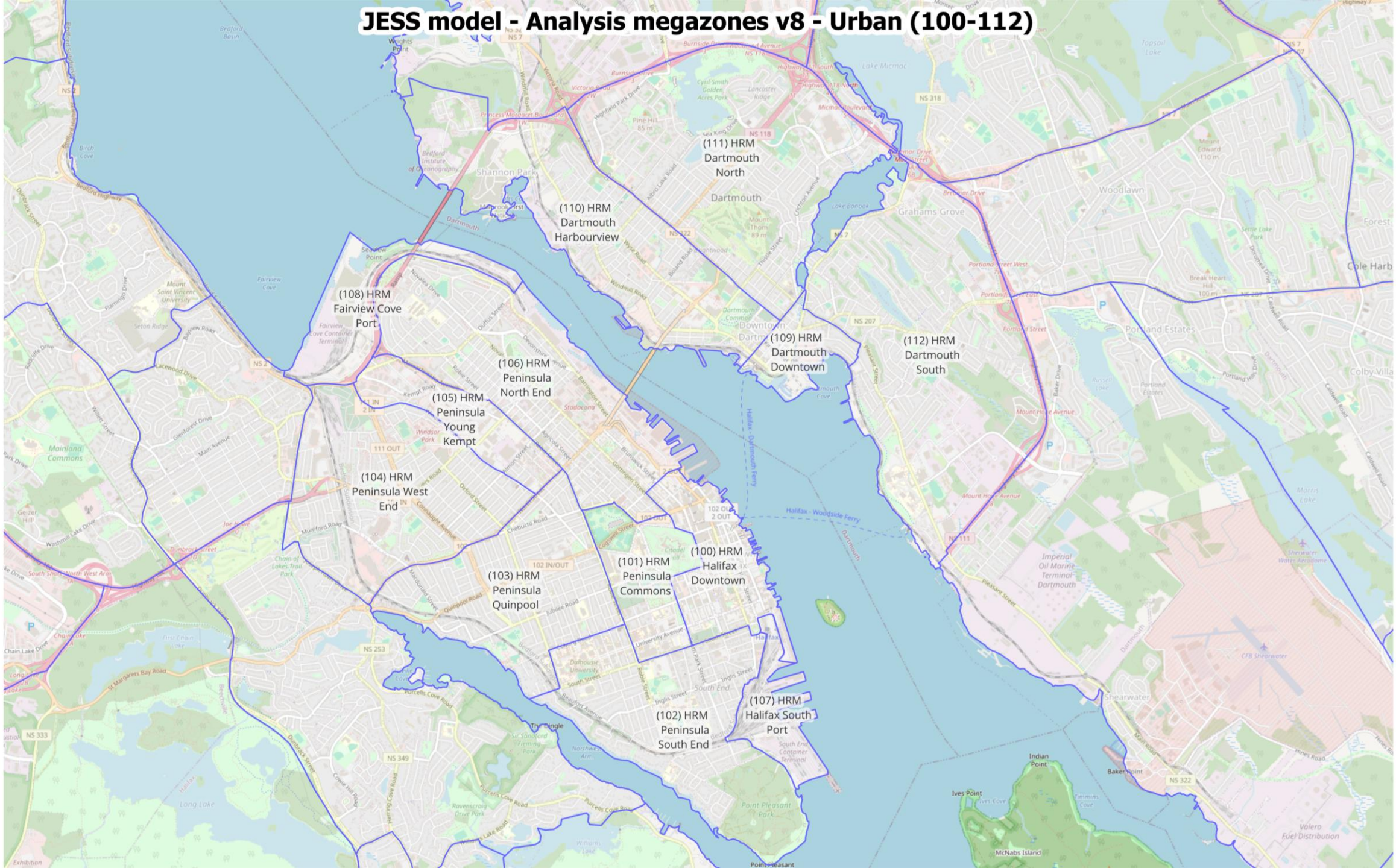
A proposal for Place of Residence-Place of Work model calibration

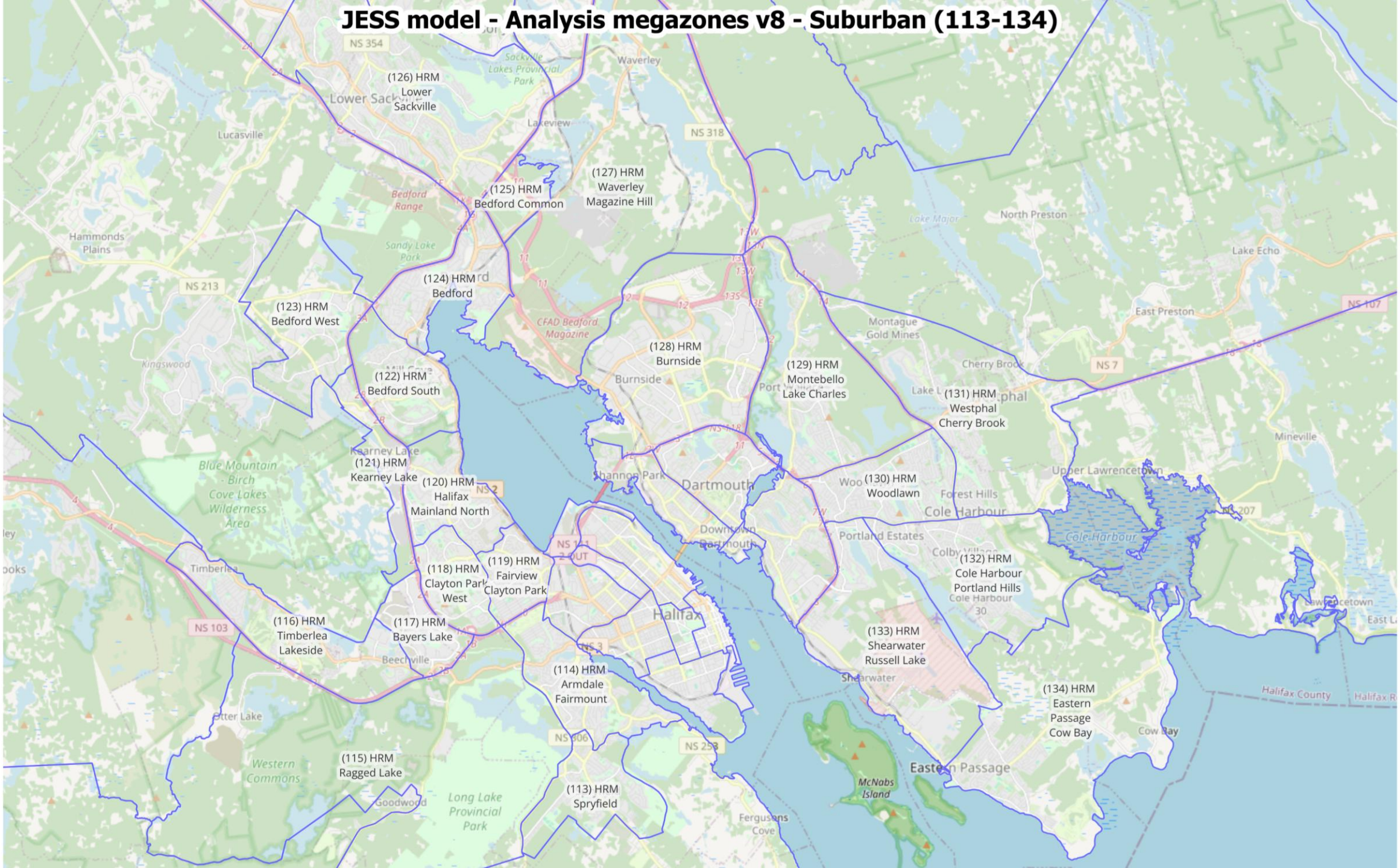
Stephen McCarthy

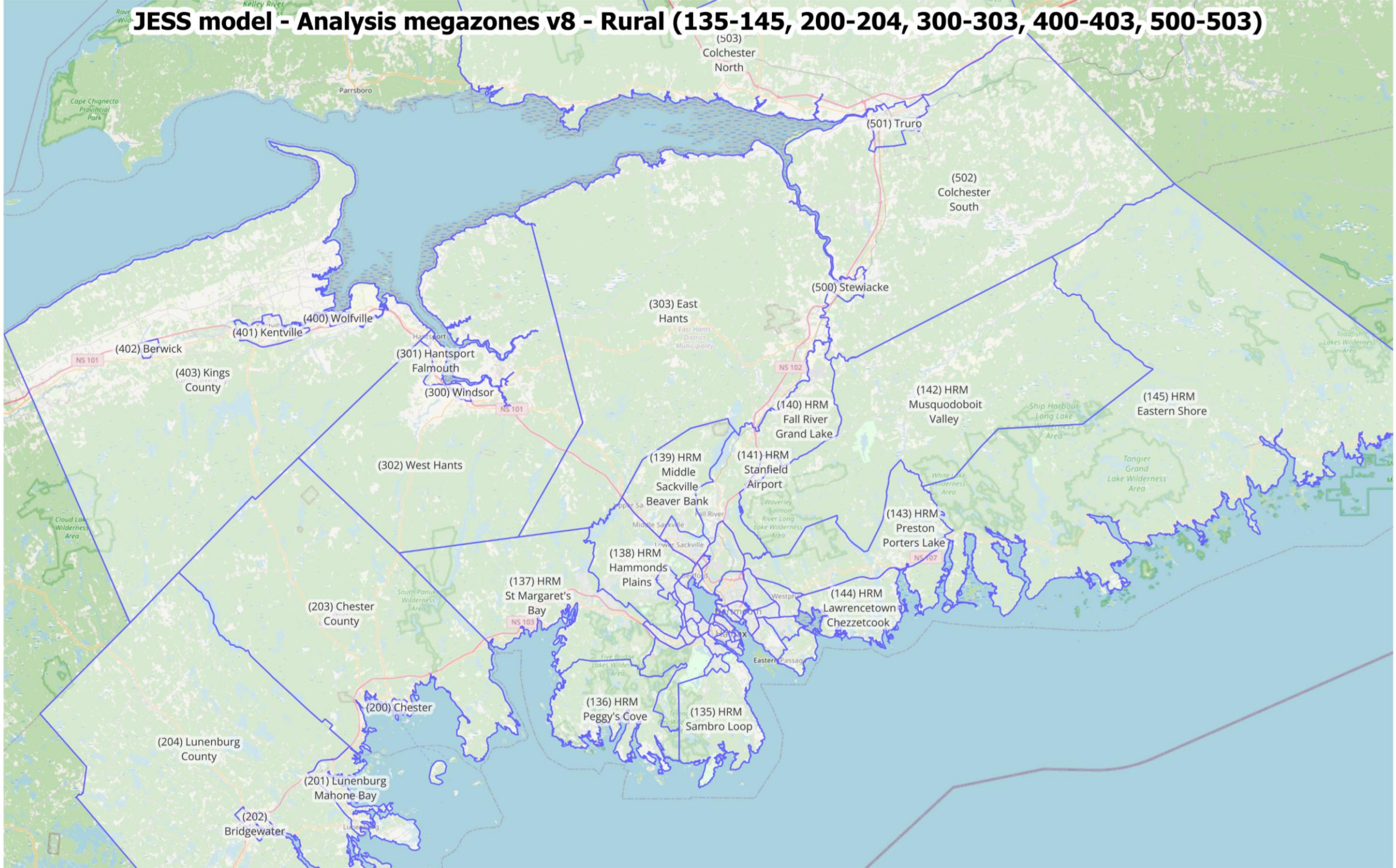
Nov 2025

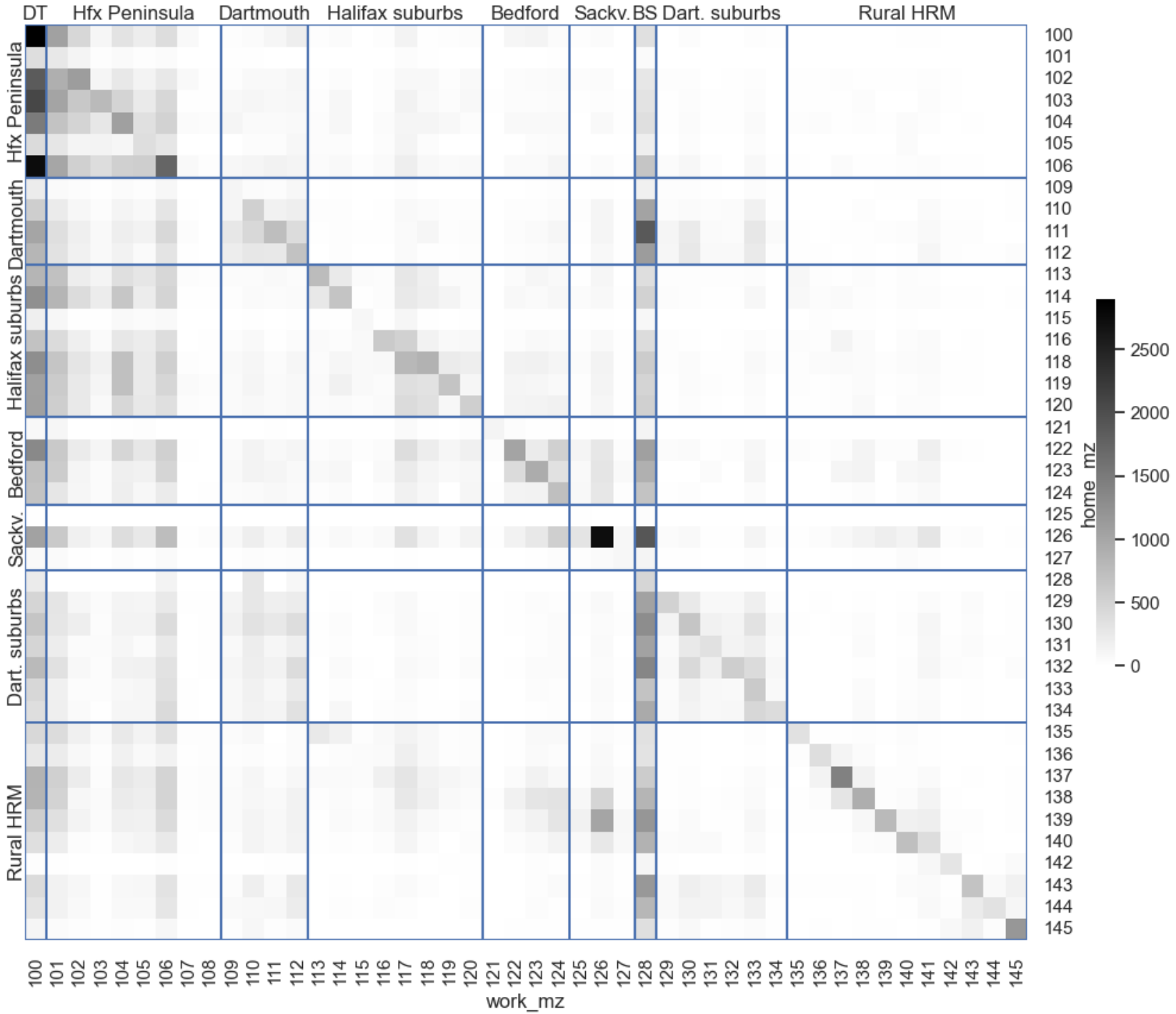
JESS model - Analysis megazones v8 - Urban (100-112)



JESS model - Analysis megazones v8 - Suburban (113-134)




JESS model - Analysis megazones v8 - Rural (135-145, 200-204, 300-303, 400-403, 500-503)
(503) Colchester North
(501) Truro
(502) Colchester South
(500) Stewiacke
(303) East Hants
(400) Wolfville
(401) Kentville
(402) Berwick
(403) Kings County
(301) Hantsport Falmouth
(300) Windsor
(302) West Hants
(140) HRM Fall River Grand Lake
(142) HRM Musquodoboit Valley
(145) HRM Eastern Shore
(139) HRM Middle Sackville Beaver Bank
(141) HRM Stanfield Airport
(143) HRM Preston Porters Lake
(138) HRM Hammonds Plains
(137) HRM St Margaret's Bay
(144) HRM Lawrencetown Chezzetcook
(203) Chester County
(200) Chester
(136) HRM Peggy's Cove
(135) HRM Sambro Loop
(204) Lunenburg County
(201) Lunenburg Mahone Bay
(202) Bridgewater



Commute O/D matrix: observed*

* From combined Census 2016 and 2021 data

Place of Residence – Place of Work (PoRPoW) gravity model: friction

$$e^{Friction_{ijk}} = K_{ij}e^{V_{ijk}}$$

Where,

$$e^{V_{ijk}} = e^{\beta_{Constant_s}} \left(\begin{cases} e^{\beta_{intrazonal_s}} & i = j \\ 1 & \text{else} \end{cases} \right) \left(\begin{cases} e^{\beta_{intraPD_s}} & PD_i = PD_j \\ 1 & \text{else} \end{cases} \right) \left(e^{\beta_{aivtt_{sk}} AIVTT_{ij}} + e^{\beta_{TransitConstant} + \beta_{ptivtt_s} PTIVTT_{ij}} + e^{\beta_{DitanceConstant} + \beta_{Distance} Distance_{ij}} \right)$$

Where,

$i =$ Origin Zone

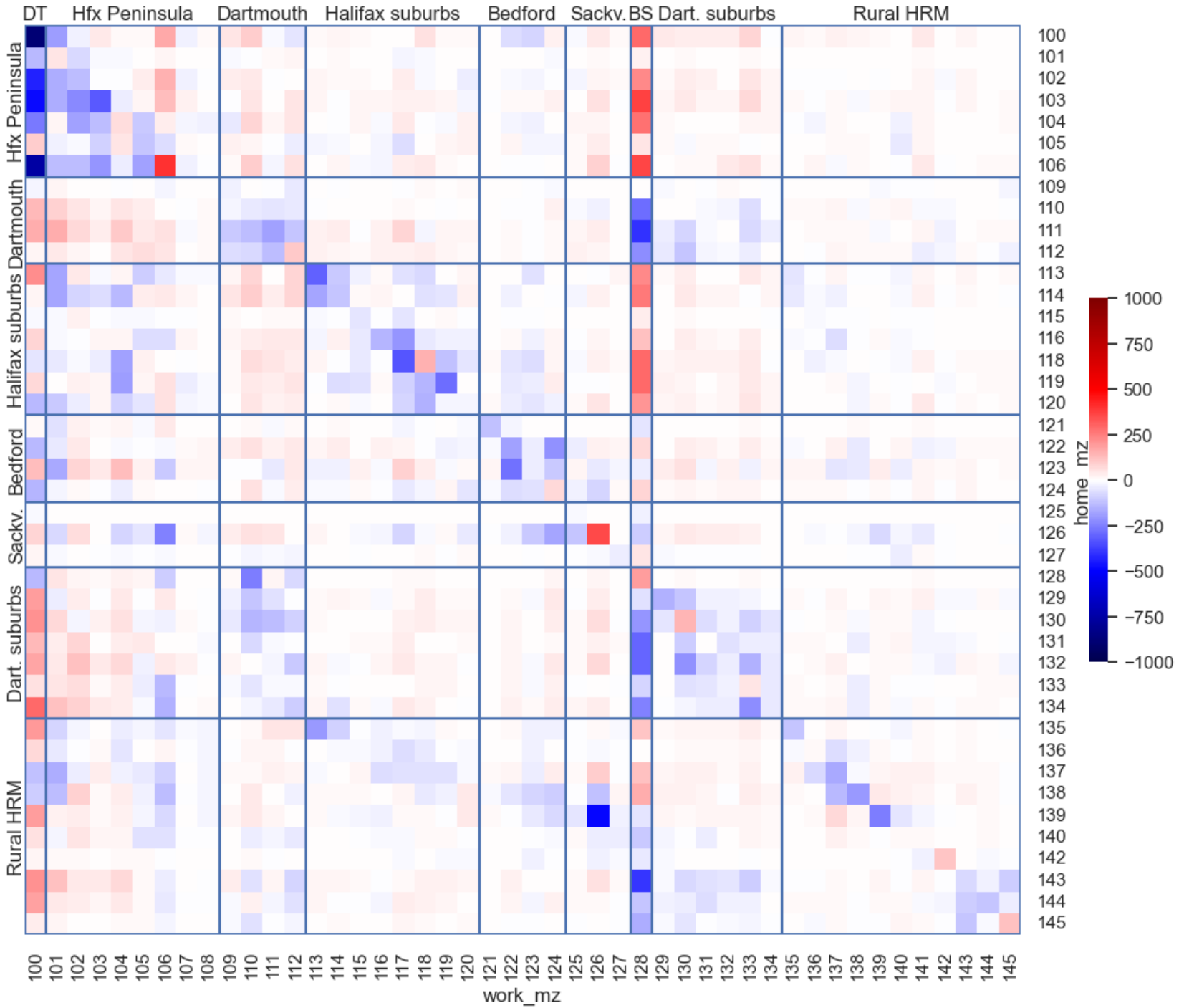
$j =$ Destination Zone

$s =$ Spatial Segment

$PD =$ Planning District

$k =$ Worker Category

$K =$ K-Factor for zone i to zone$j$



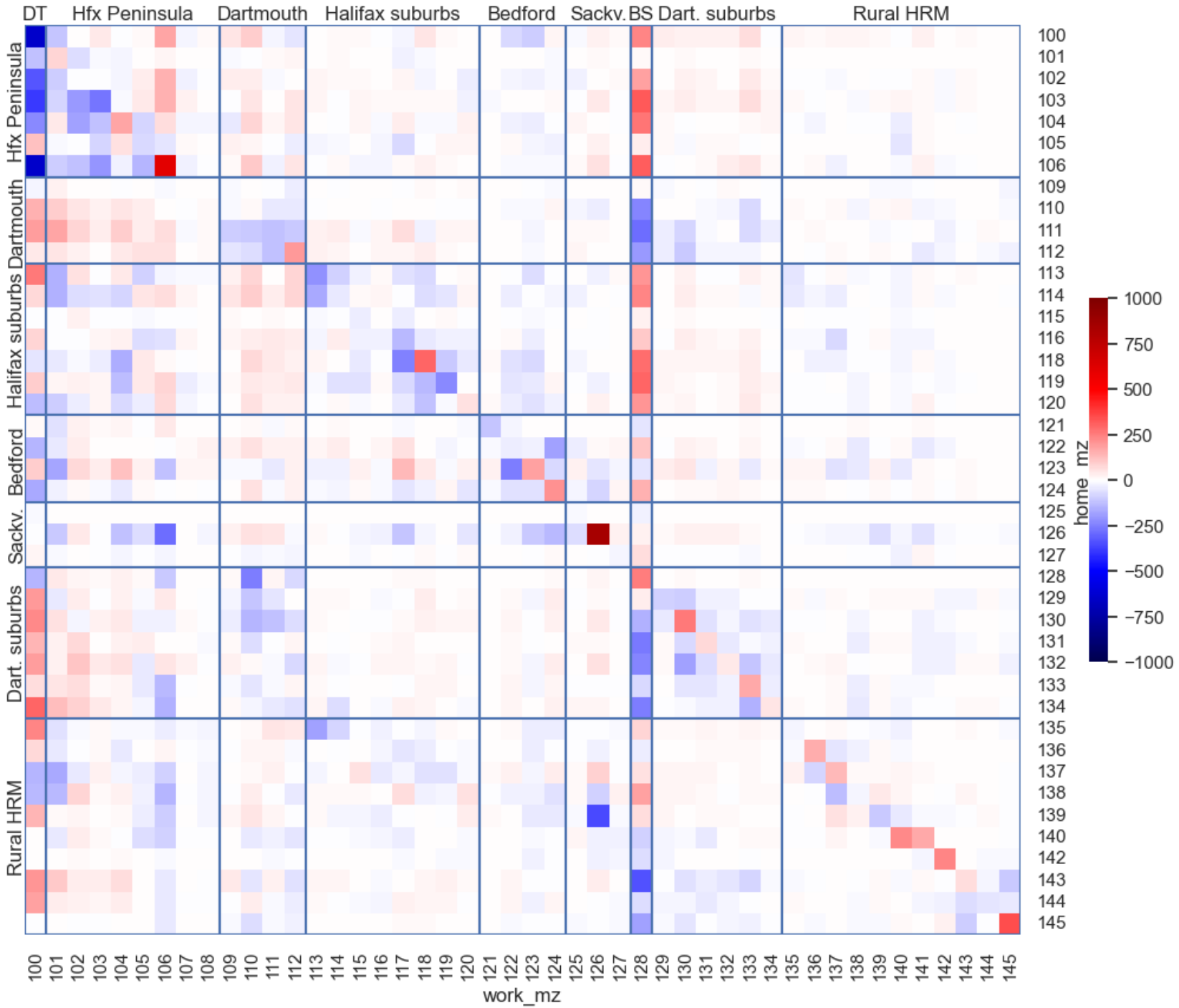
Difference in predicted flows, model vs. observed

Uncalibrated model (Toronto parameters)

[Signed] GEH statistic for difference in predicted flows, model vs. observed

Uncalibrated model (Toronto parameters)

$$GEH = \sqrt{\frac{2(M-C)^2}{M+C}}$$



Difference in predicted flows, model vs. observed

Model with 1.5x time/distance parameters

[Signed] GEH statistic for difference in predicted flows, model vs. observed

Model with 1.5x time/distance parameters

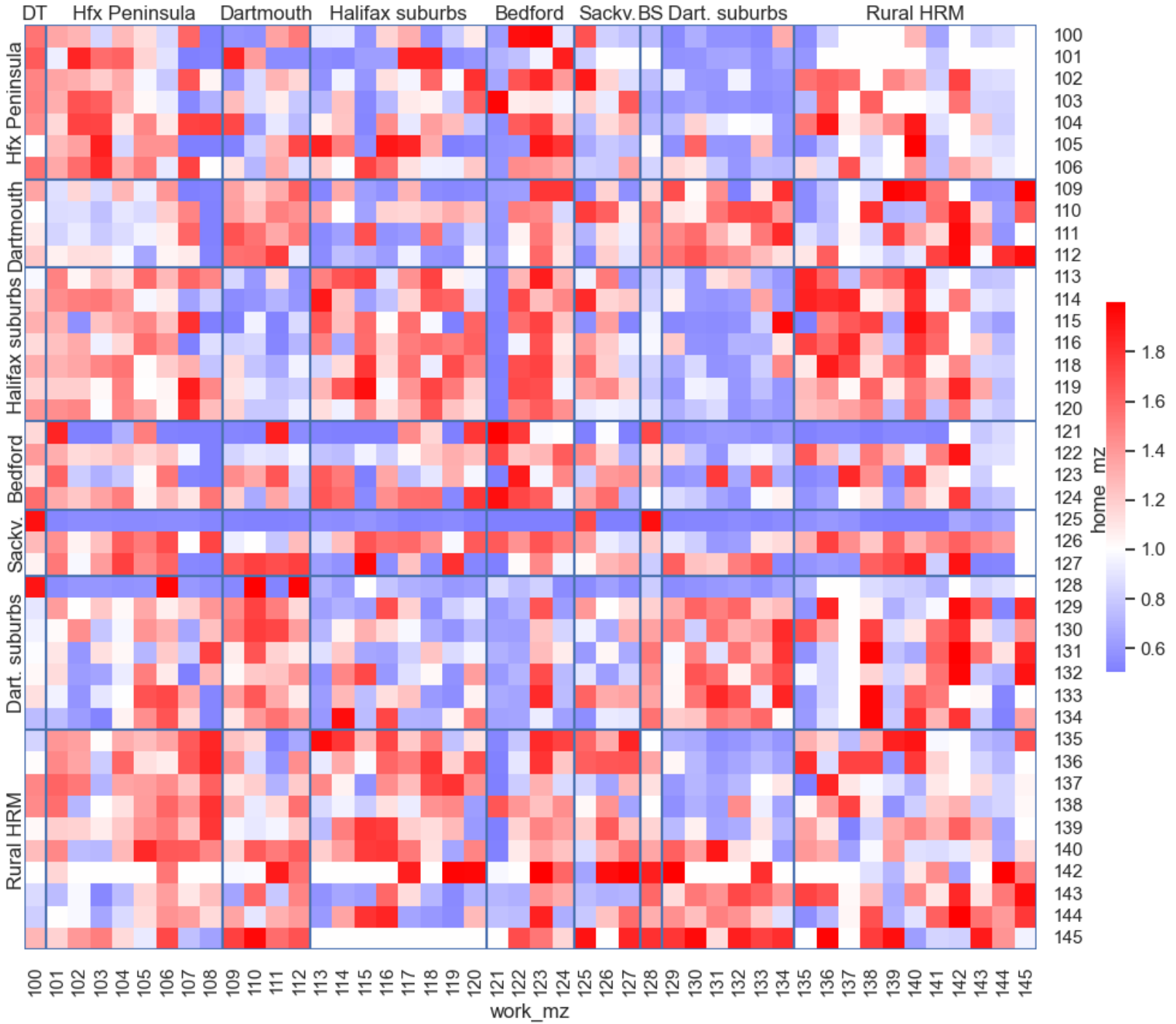
$$GEH = \sqrt{\frac{2(M-C)^2}{M+C}}$$

Question: how to calibrate

- Current approach
  - K factors: manually choose pairs of O and D ranges and value to multiply friction matrix

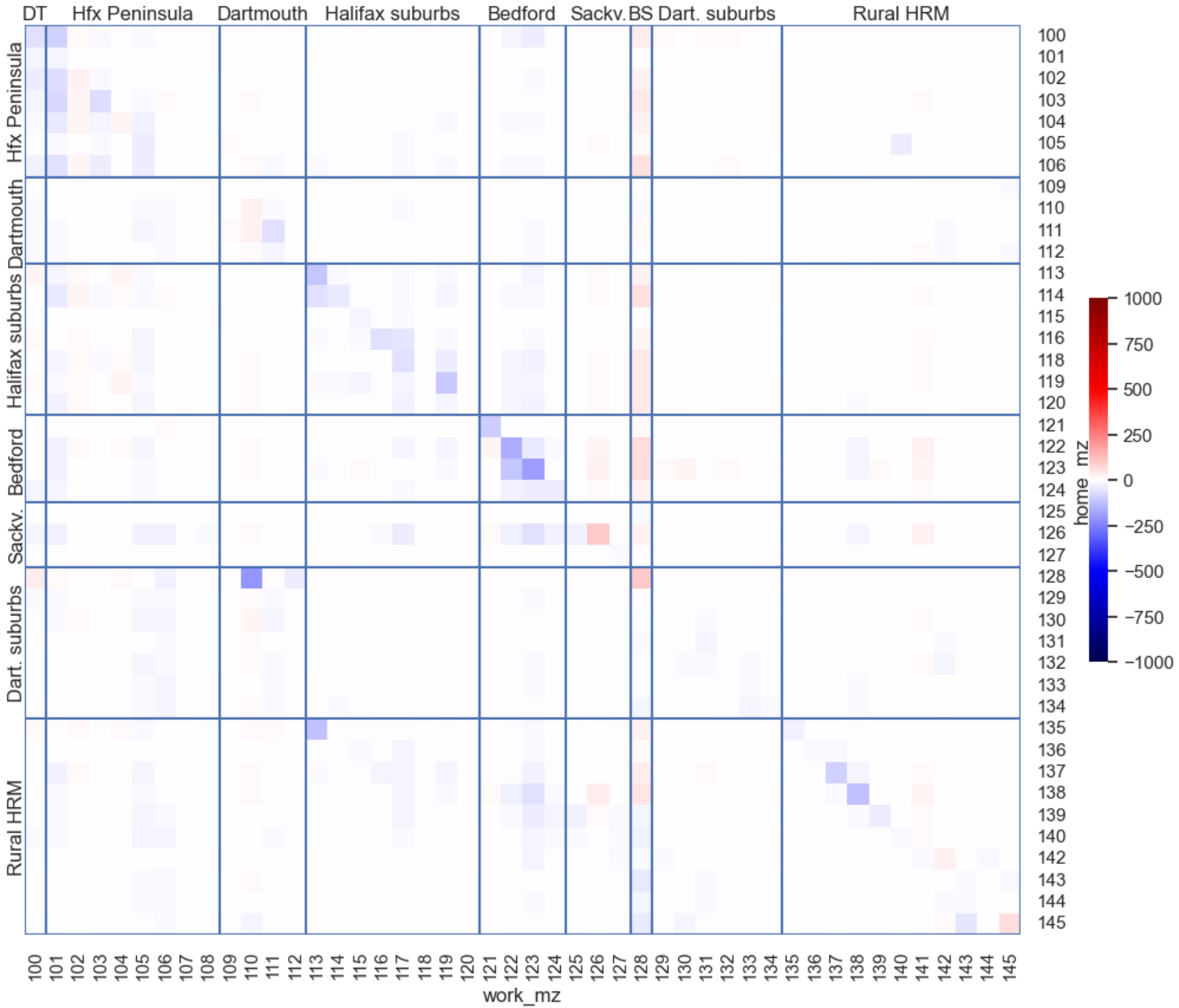
$$e^{Friction_{ijk}} = K_{ij}e^{V_{ijk}}$$

- Desired approach
  - Find algorithm to calculate ‘best’ K-factors across all O/D pairs



Naïve K-factor matrix (clipped to [1/2, 2])

Model with 1.5x time/distance parameters



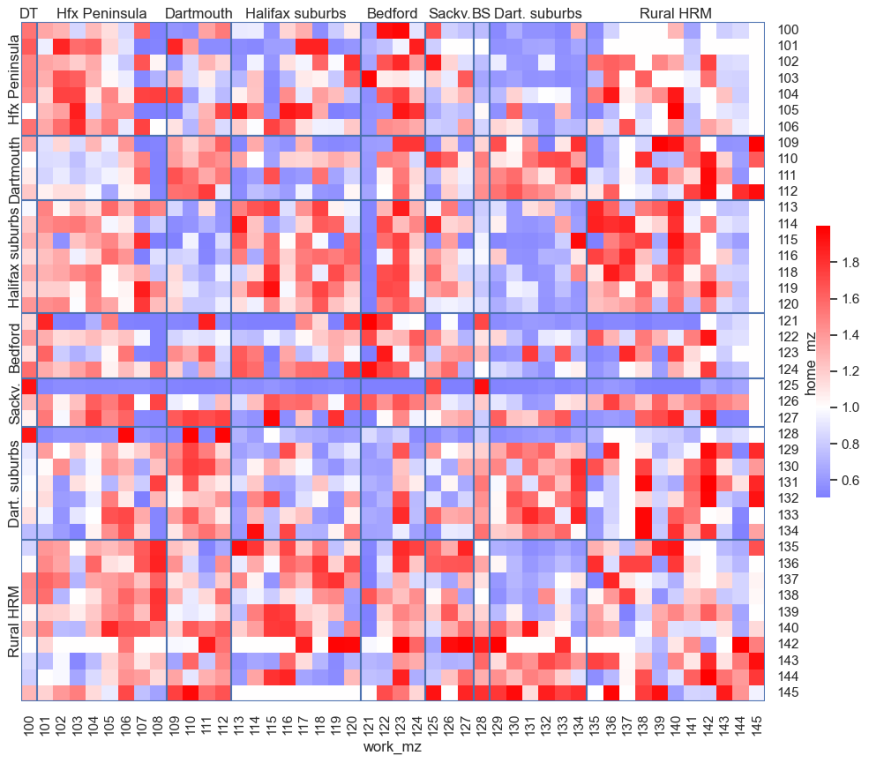
Difference in predicted flows, model vs. observed

Model with naïve K-factor calculation

Issue: overfitting!

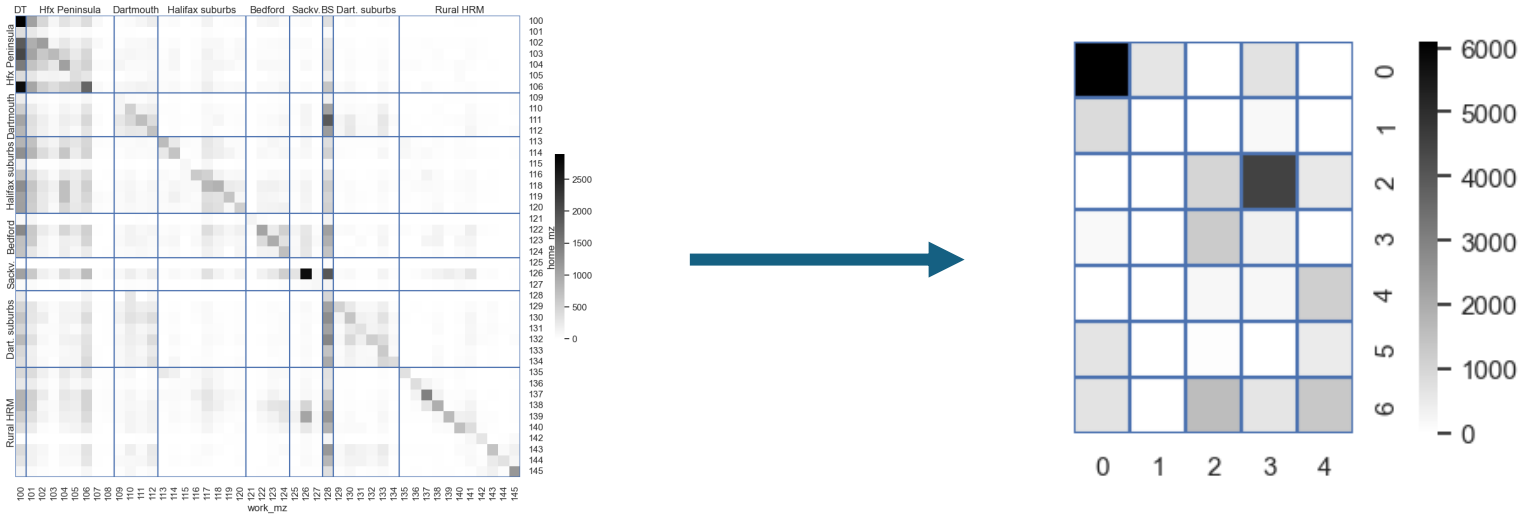
- Observed O/D data not perfect (especially in newly developed areas)
- Using the naïve approach effectively overrides the model

- Is there a way to calculate good K-factors which capture the overall patterns?

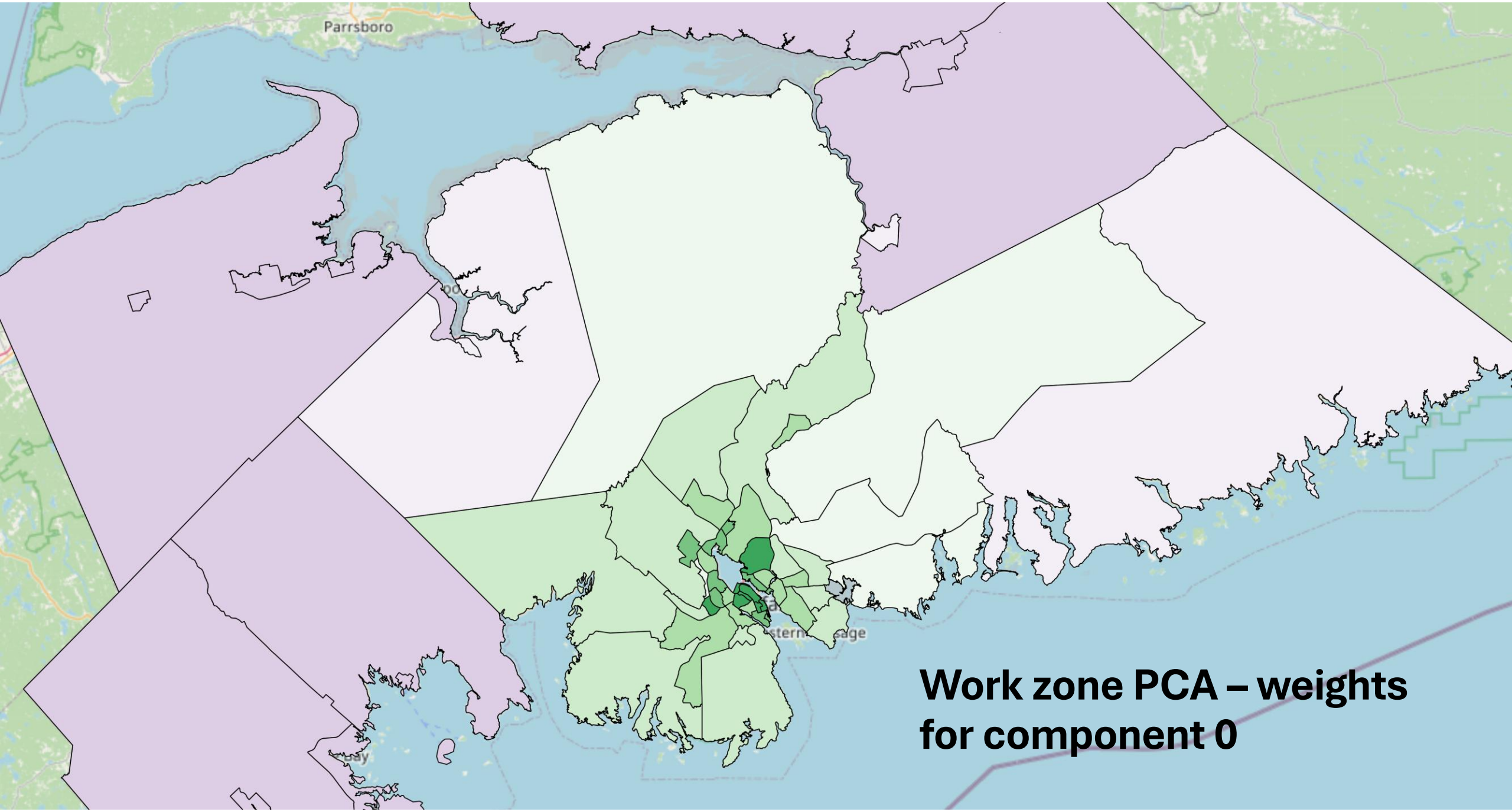


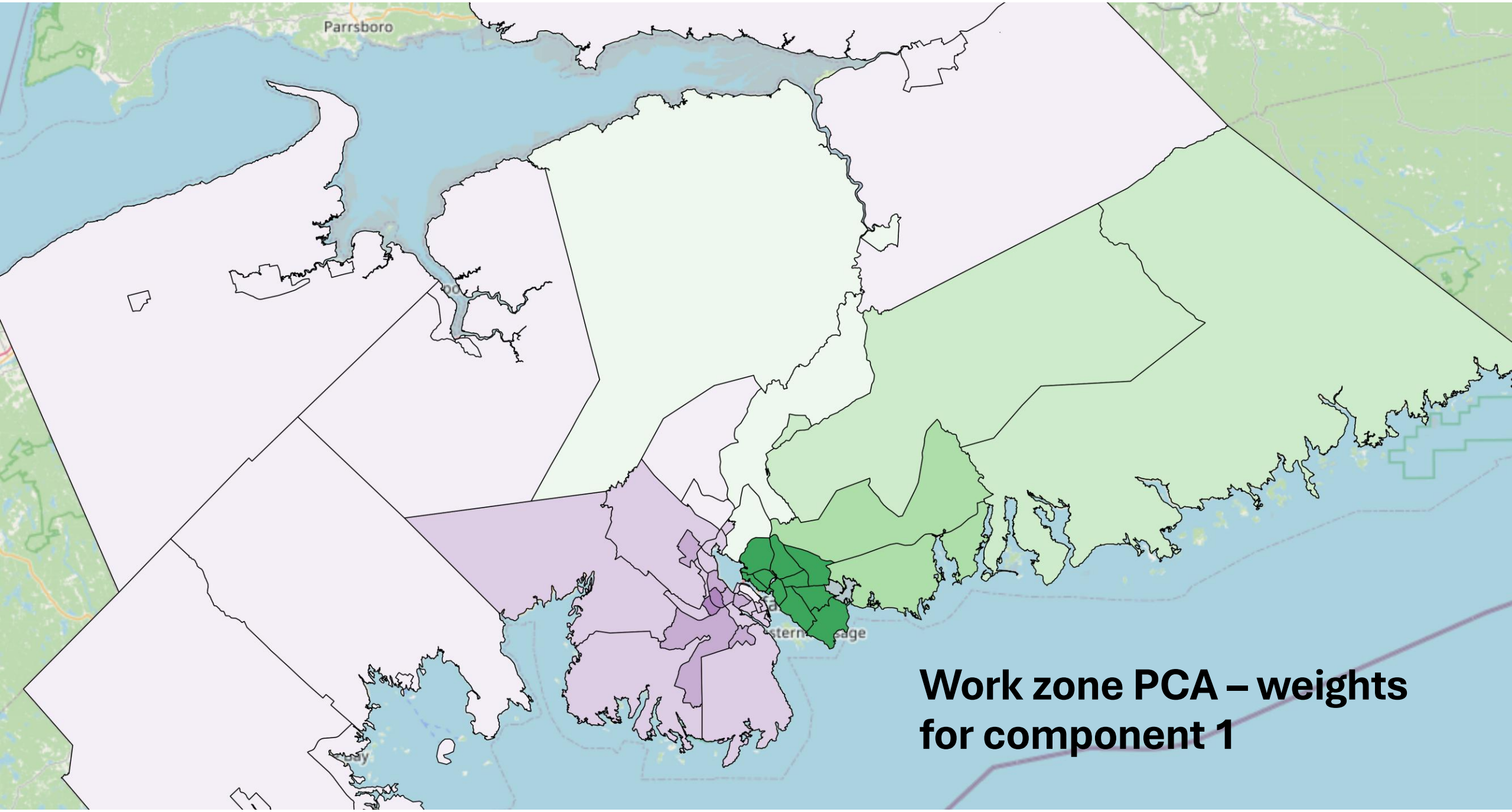
Idea: Principal component analysis

- Use PCA on the observed O/D matrix to distill the main components of variation in both dimensions

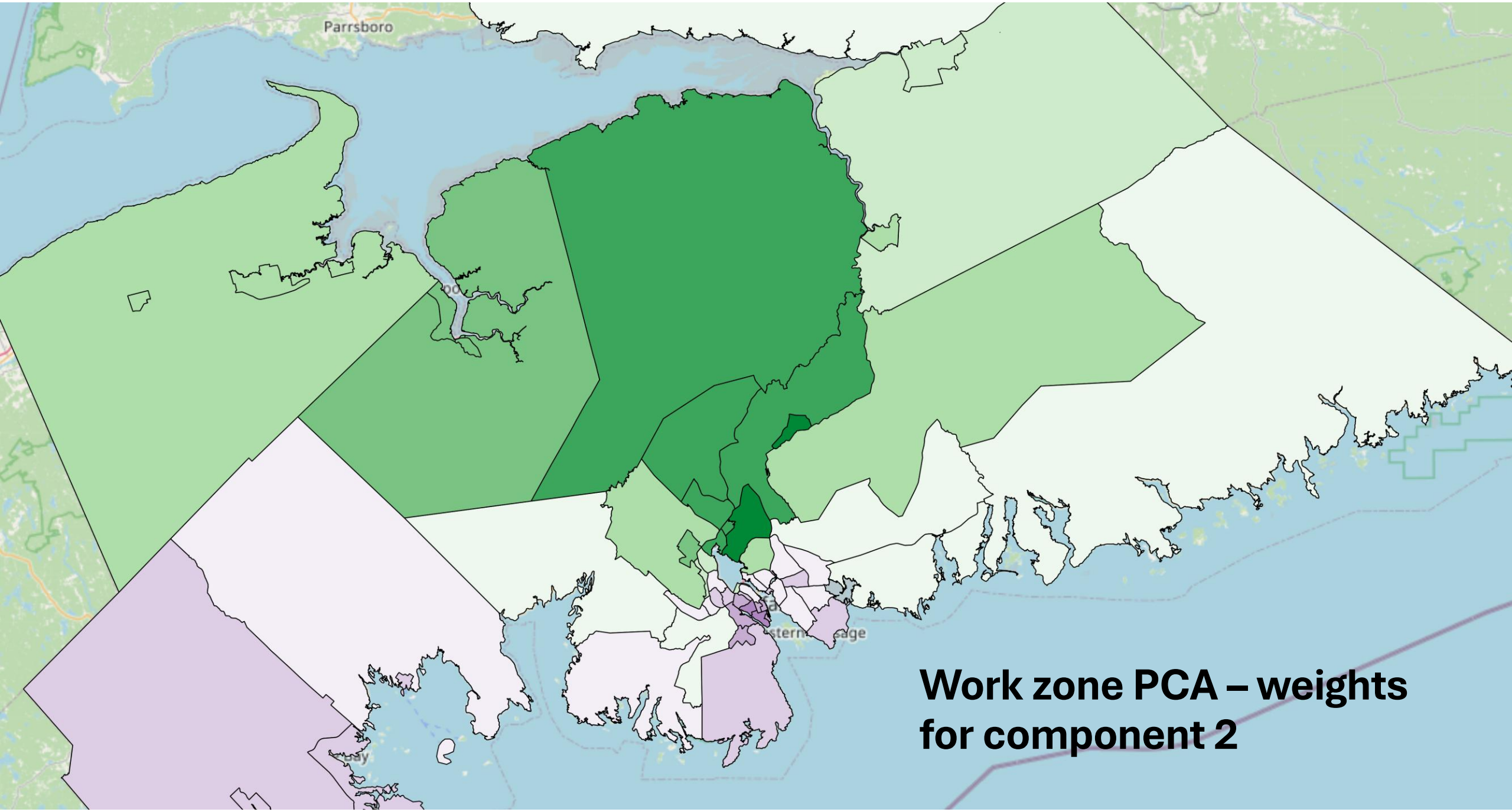


- Choose number of PCs based on how many represent reasonable geographical patterns?

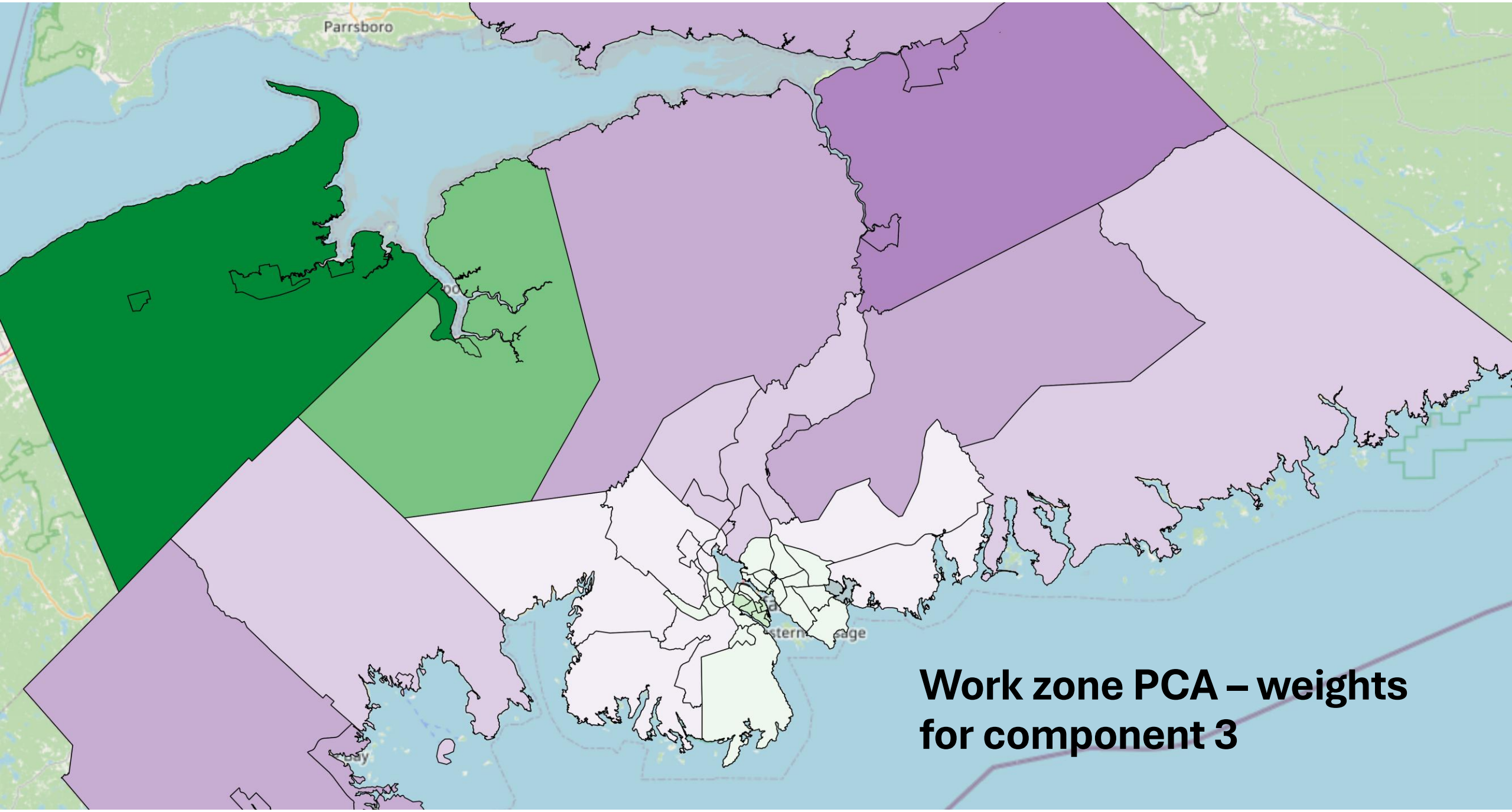

Parrsboro
Work zone PCA – weights
for component 0



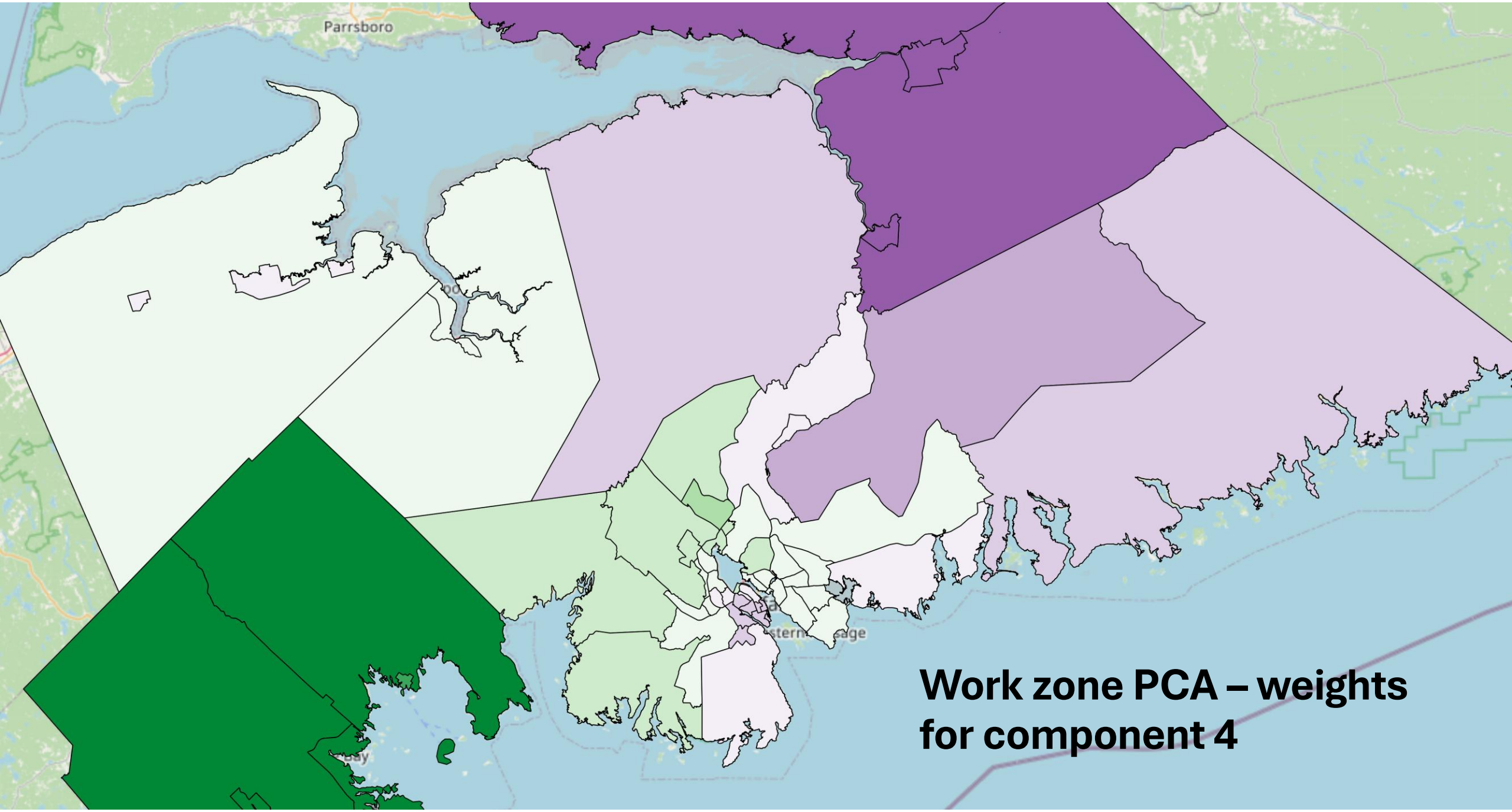
Work zone PCA – weights for component 1

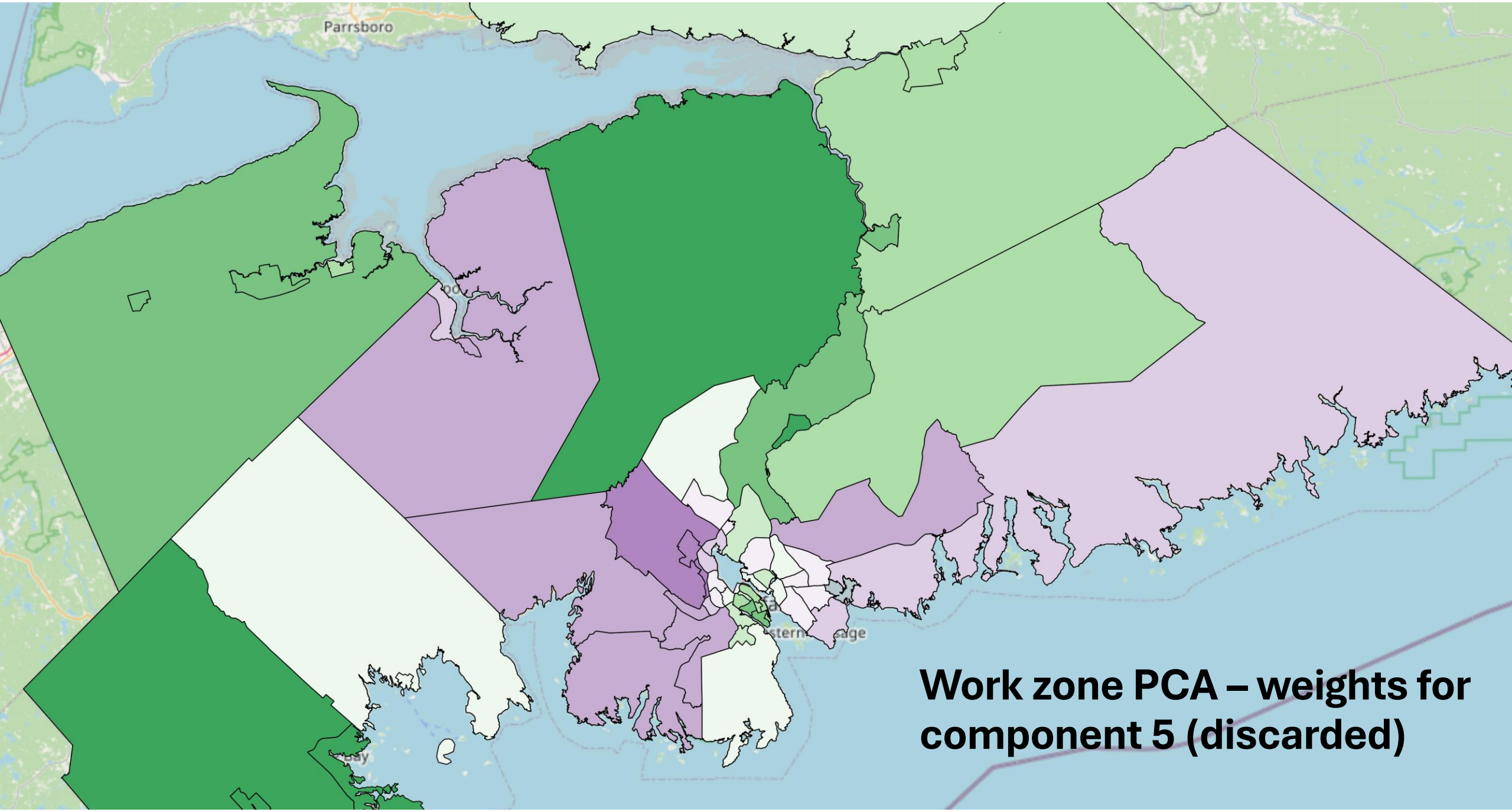


Work zone PCA – weights for component 2

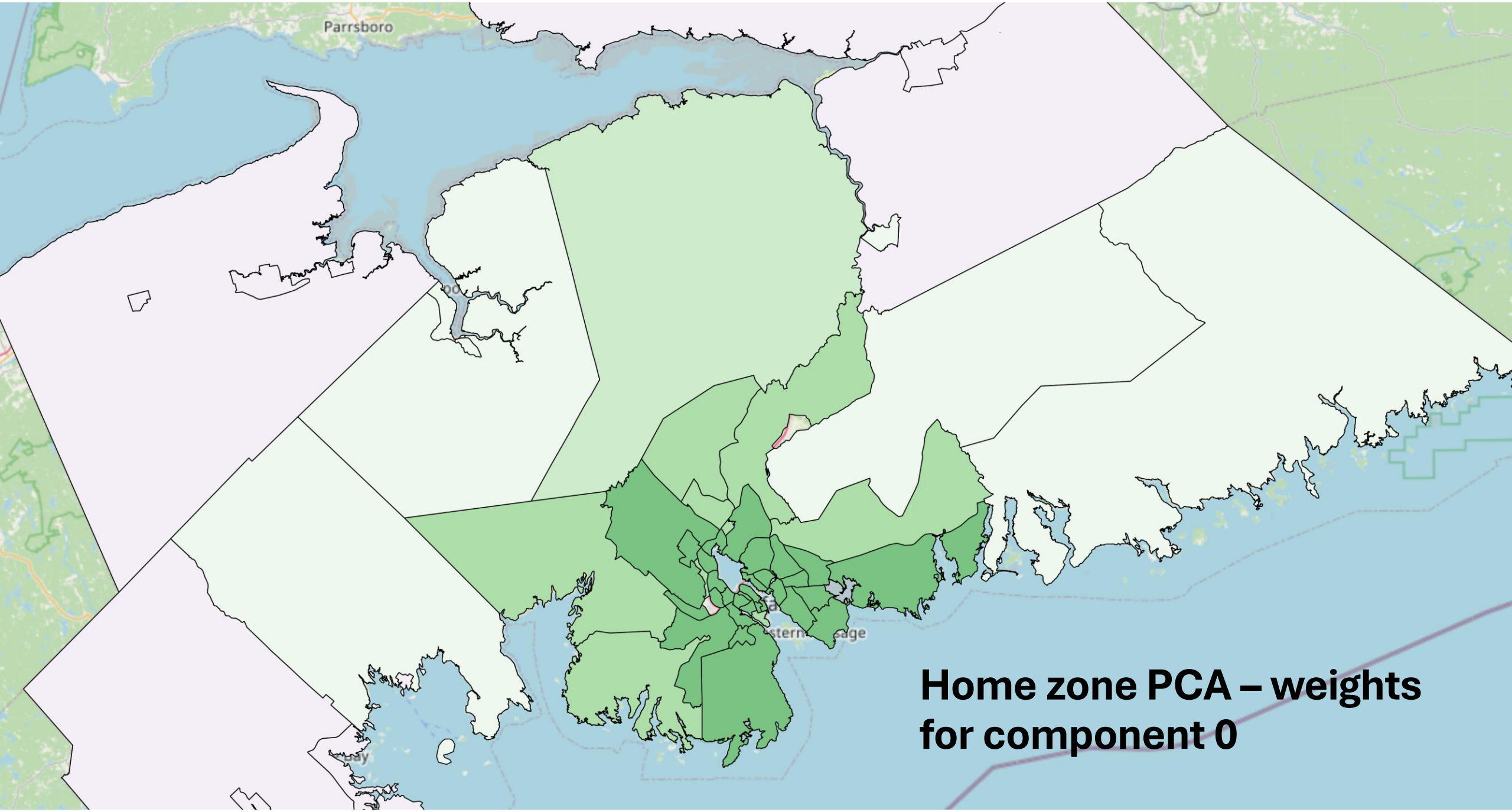


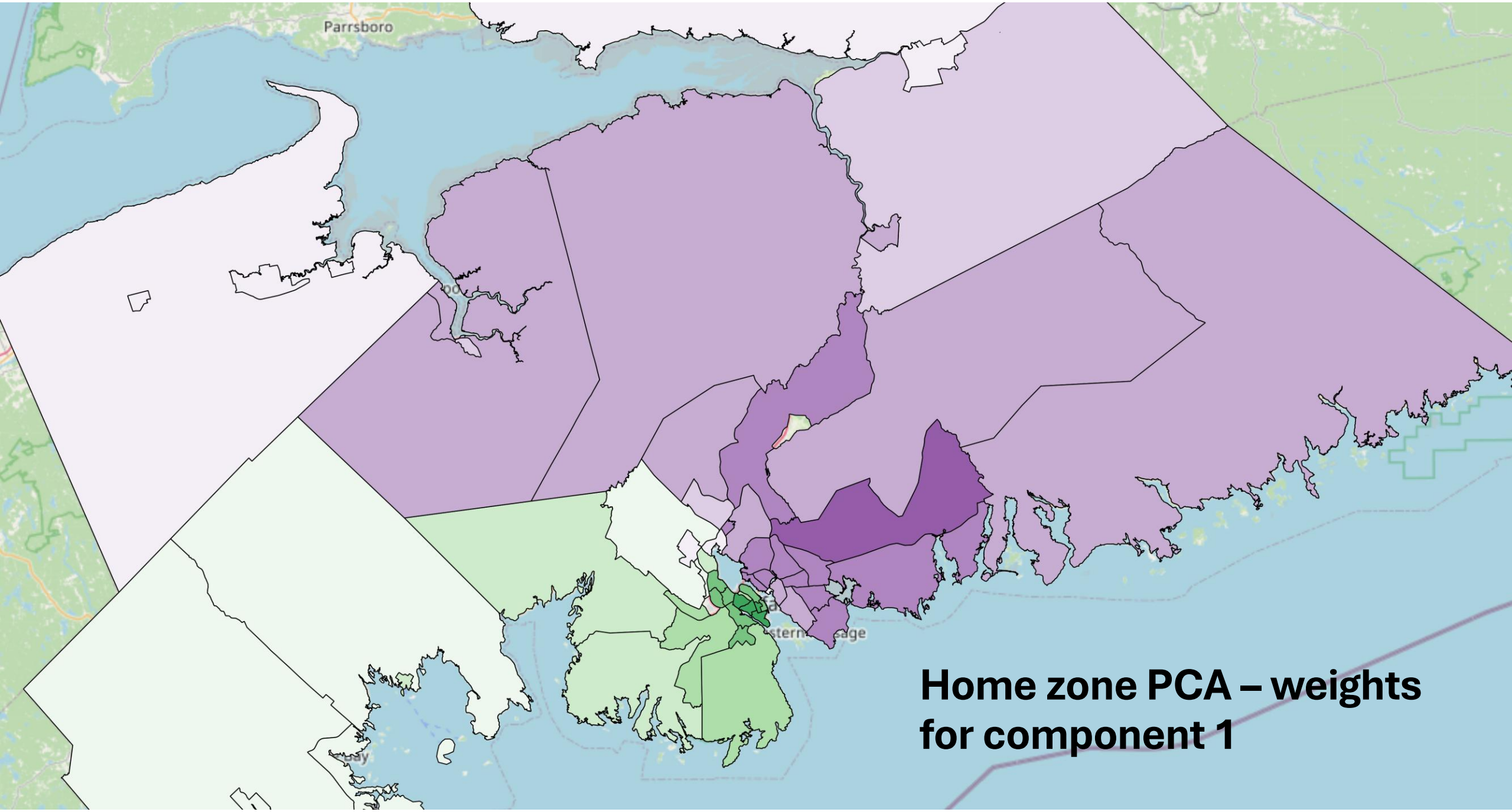
Work zone PCA – weights for component 3


Parrsboro
Work zone PCA – weights
for component 4

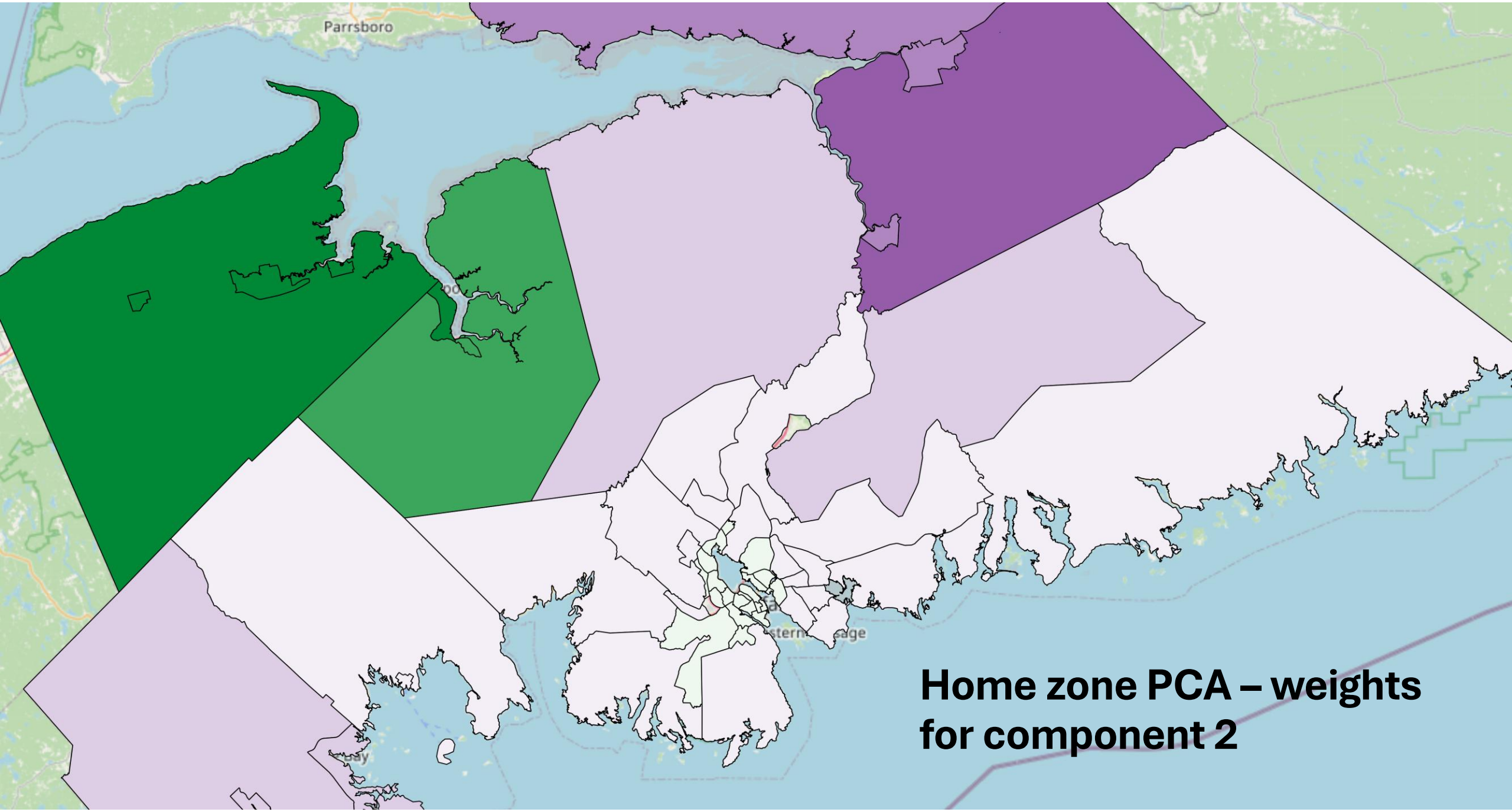


Work zone PCA – weights for component 5 (discarded)

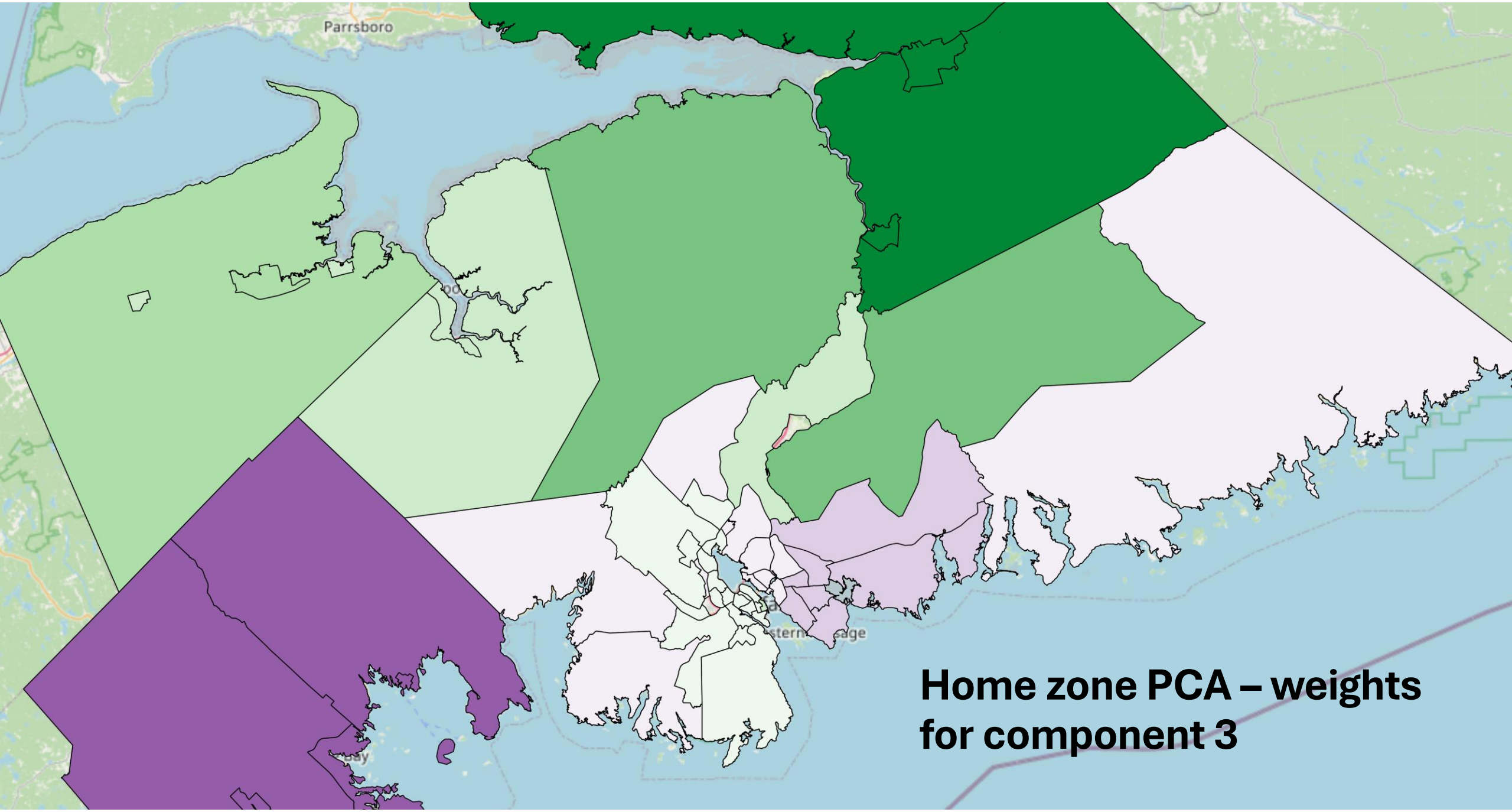

Parrsboro
Home zone PCA – weights
for component 0



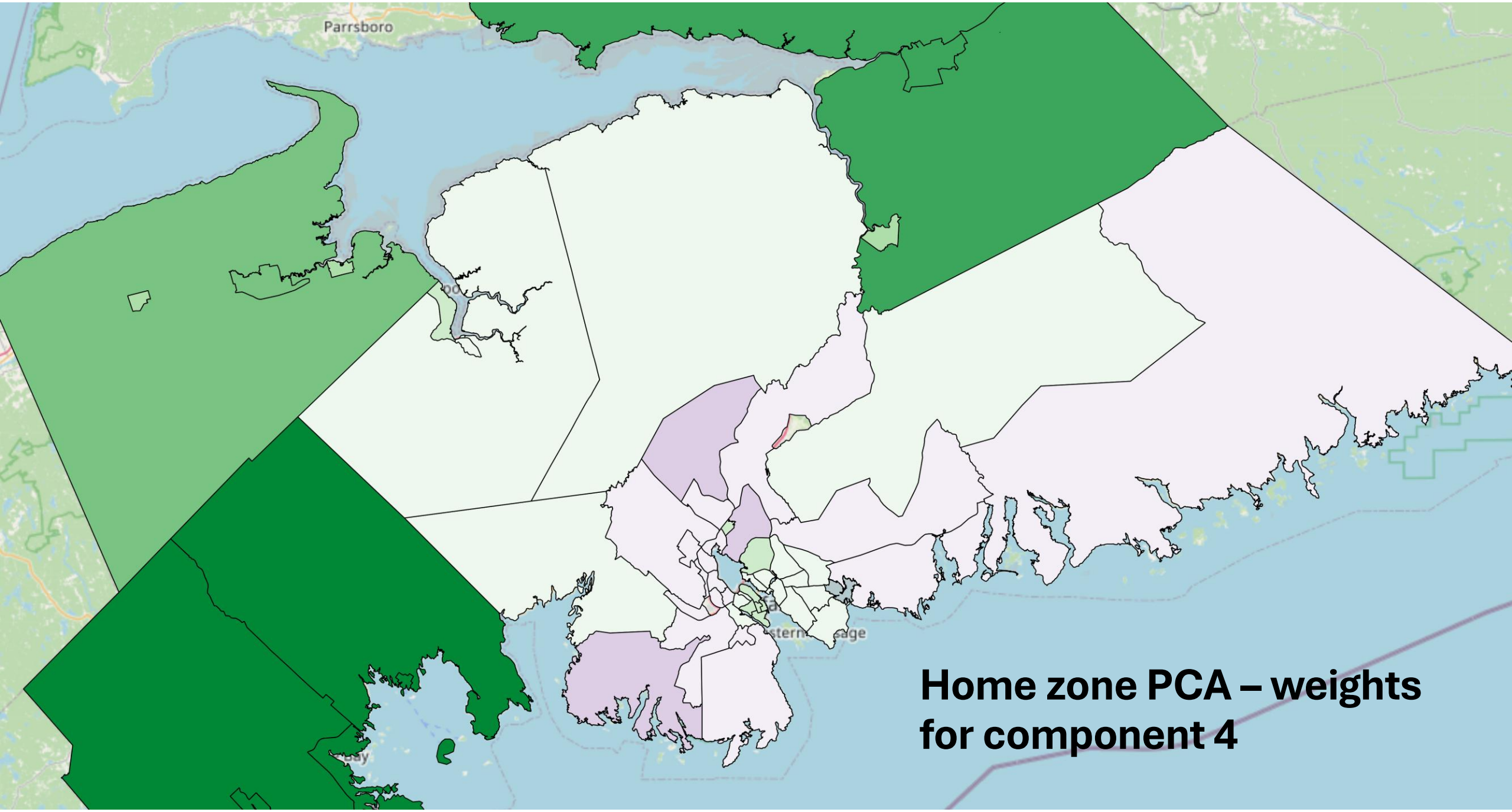
Home zone PCA – weights for component 1



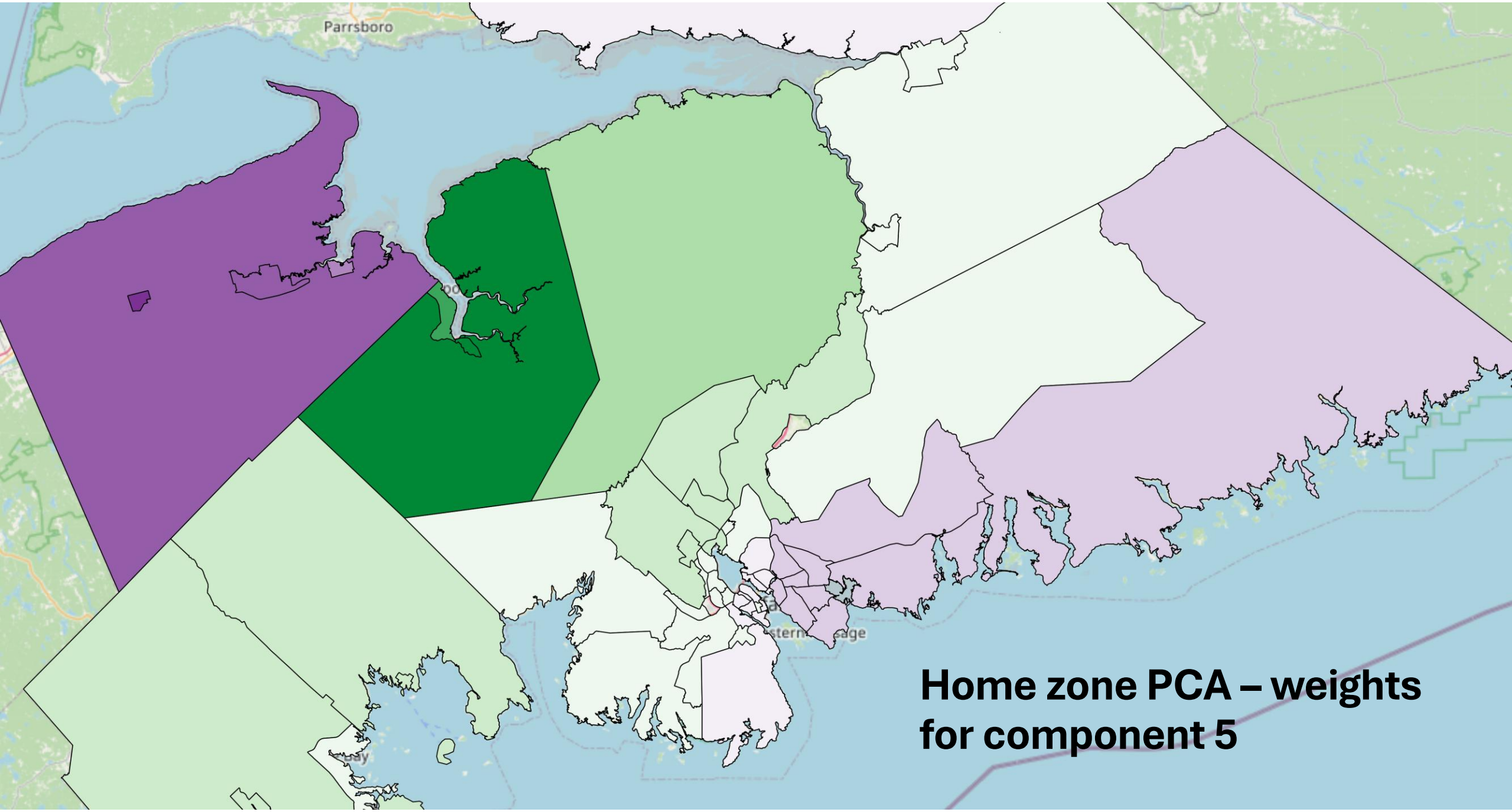
Home zone PCA – weights for component 2



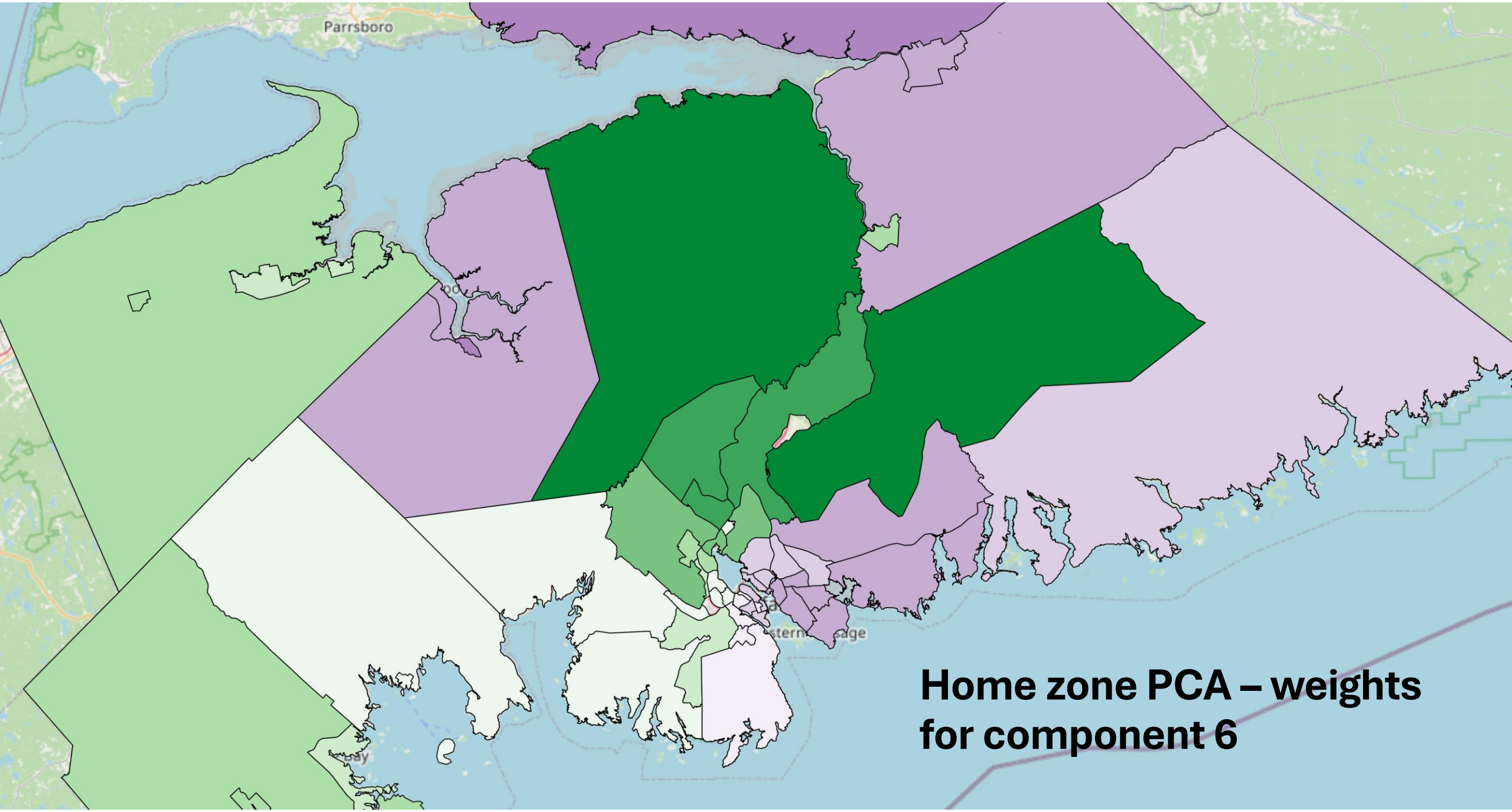
Home zone PCA – weights for component 3



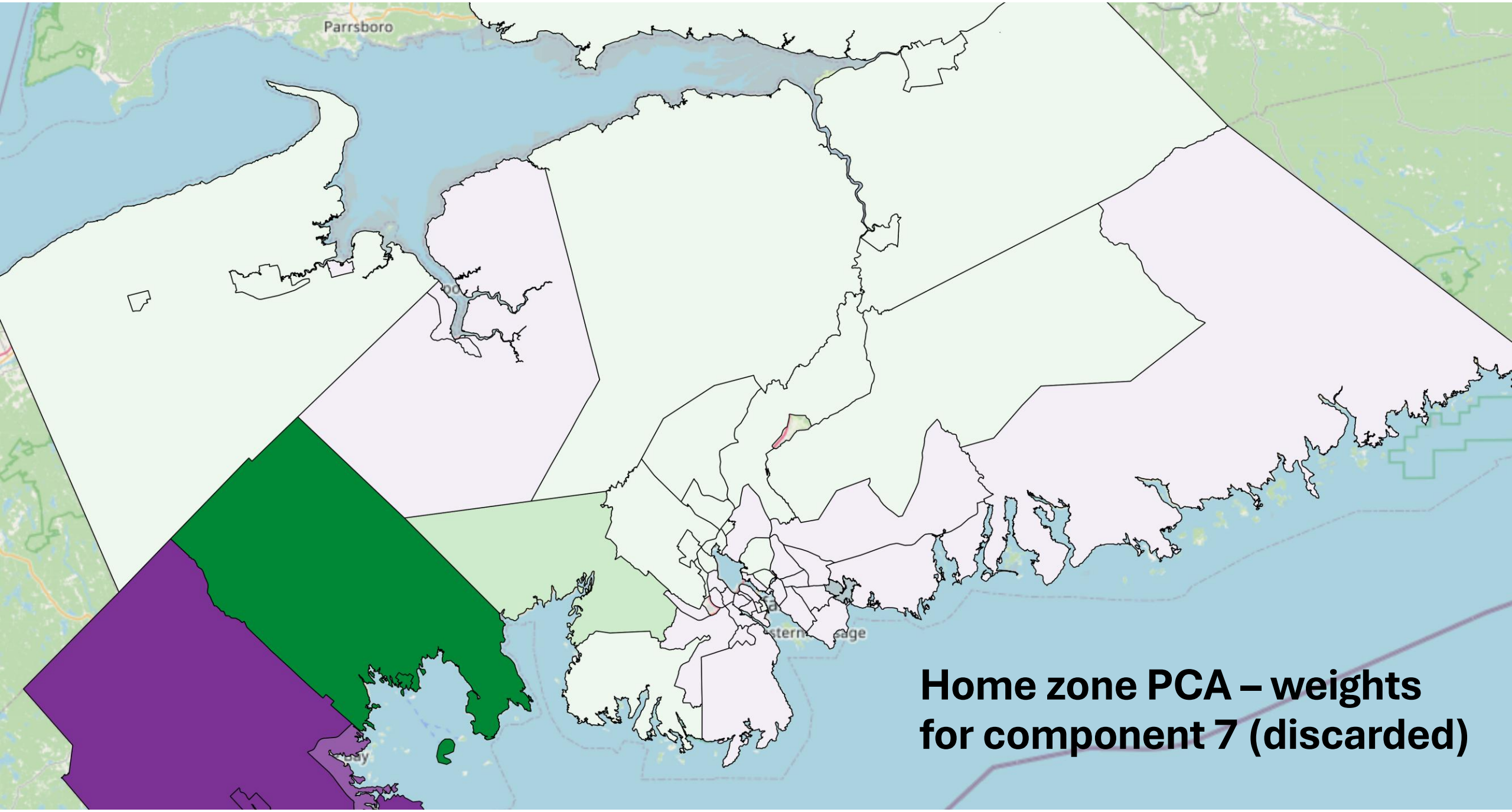
Home zone PCA – weights for component 4



Home zone PCA – weights for component 5

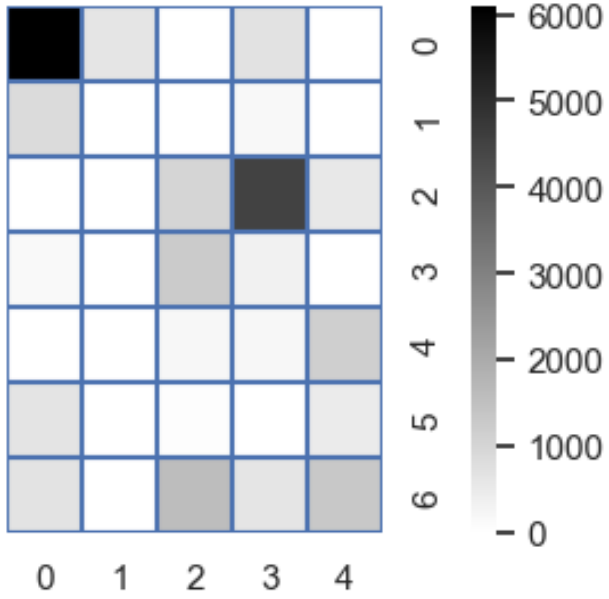


Home zone PCA – weights for component 6

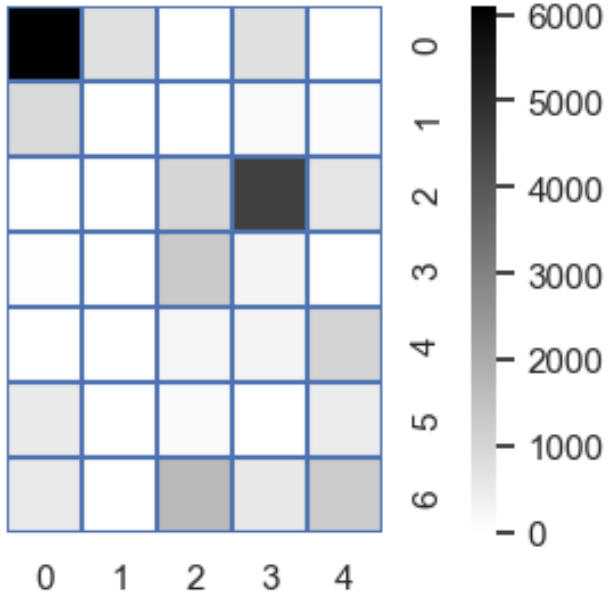

Parrsboro
Home zone PCA – weights
for component 7 (discarded)

Modelled-observed comparison on components

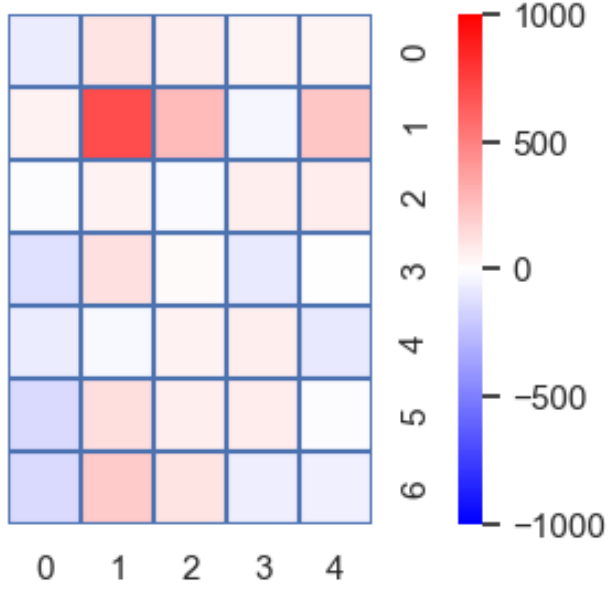
Observed



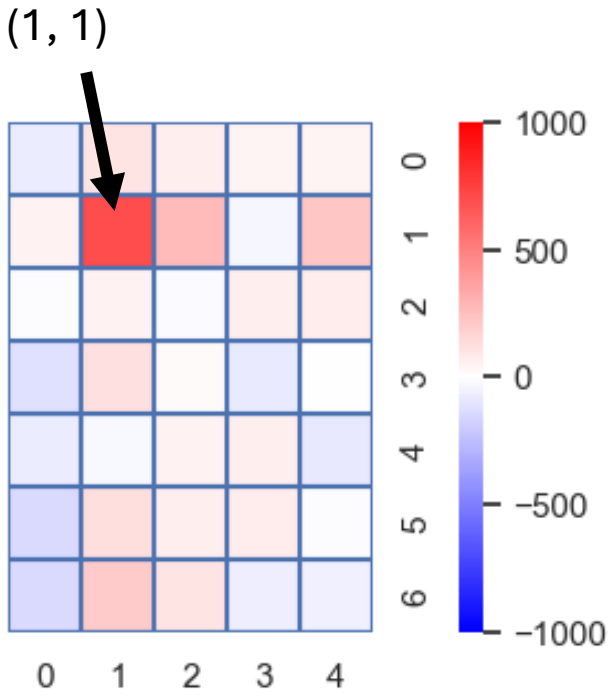
Modelled



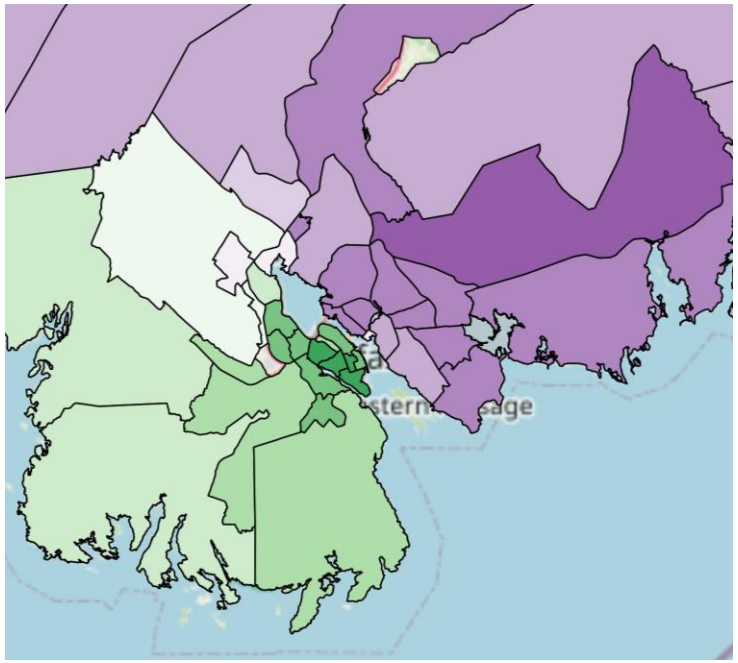
Difference (M-O)



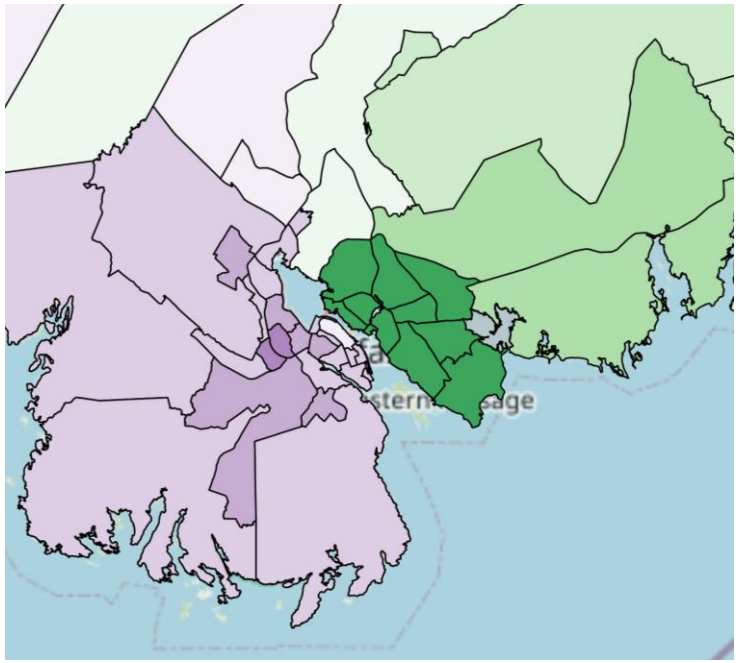
Observation about highest-difference cell



Home zone component 1



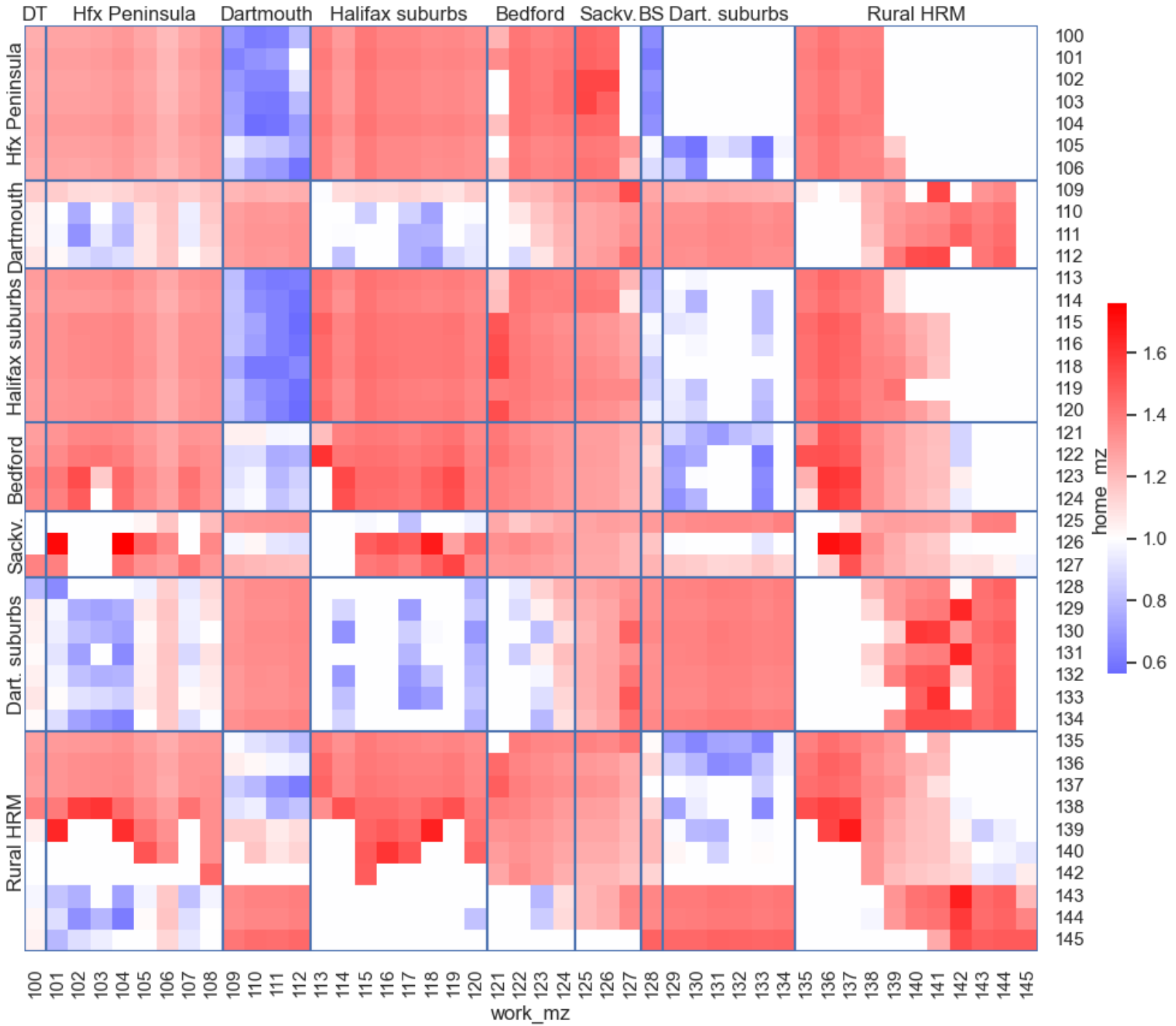
Work zone component 1



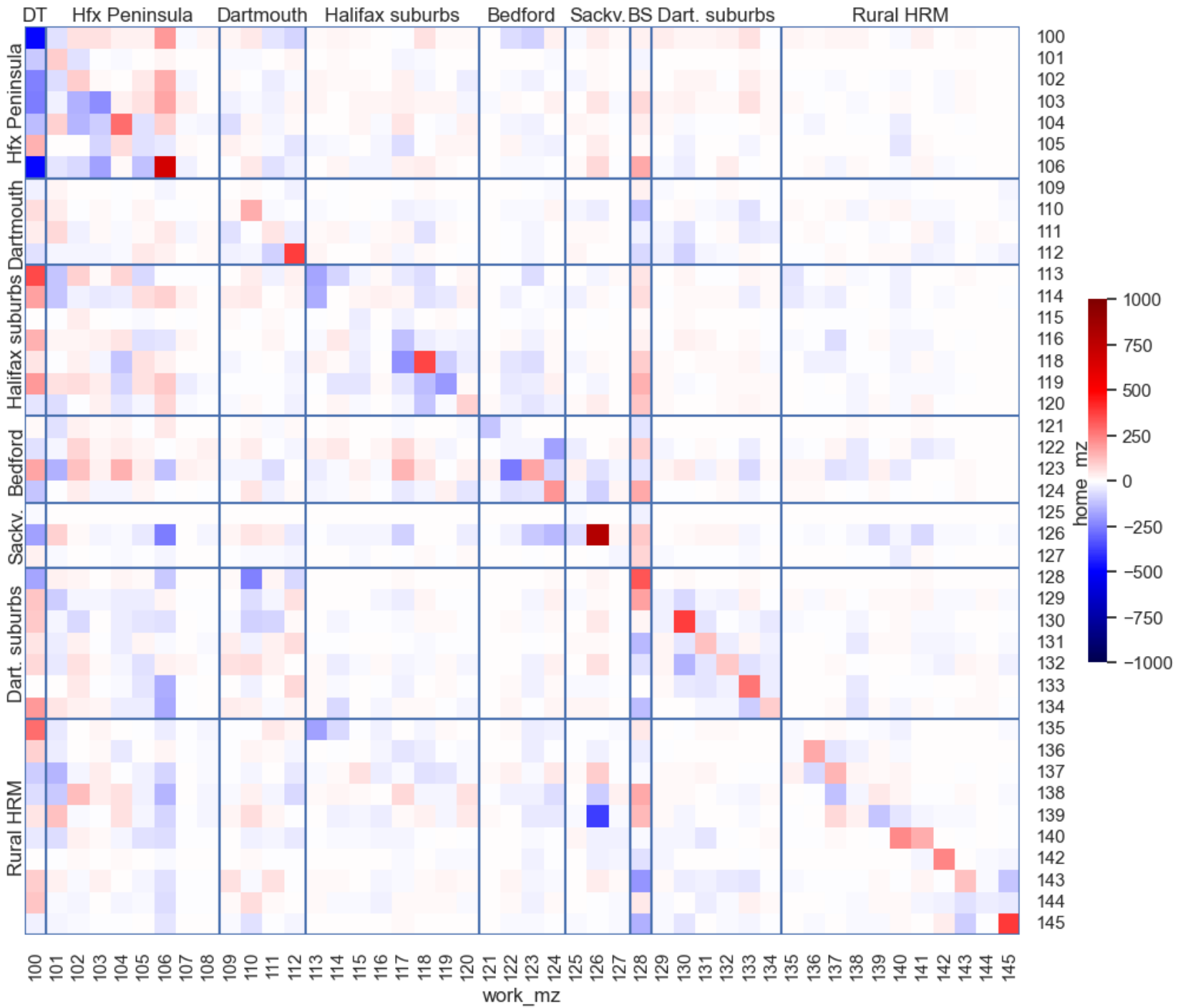
- Model overestimates cross-harbour commuting

Application to compute K-factors

- Calculate friction matrices for observed and modelled
- Reduce then re-expand both matrices in both directions to keep only the main factors
- Divide re-expanded matrices to compute K-factor matrix.

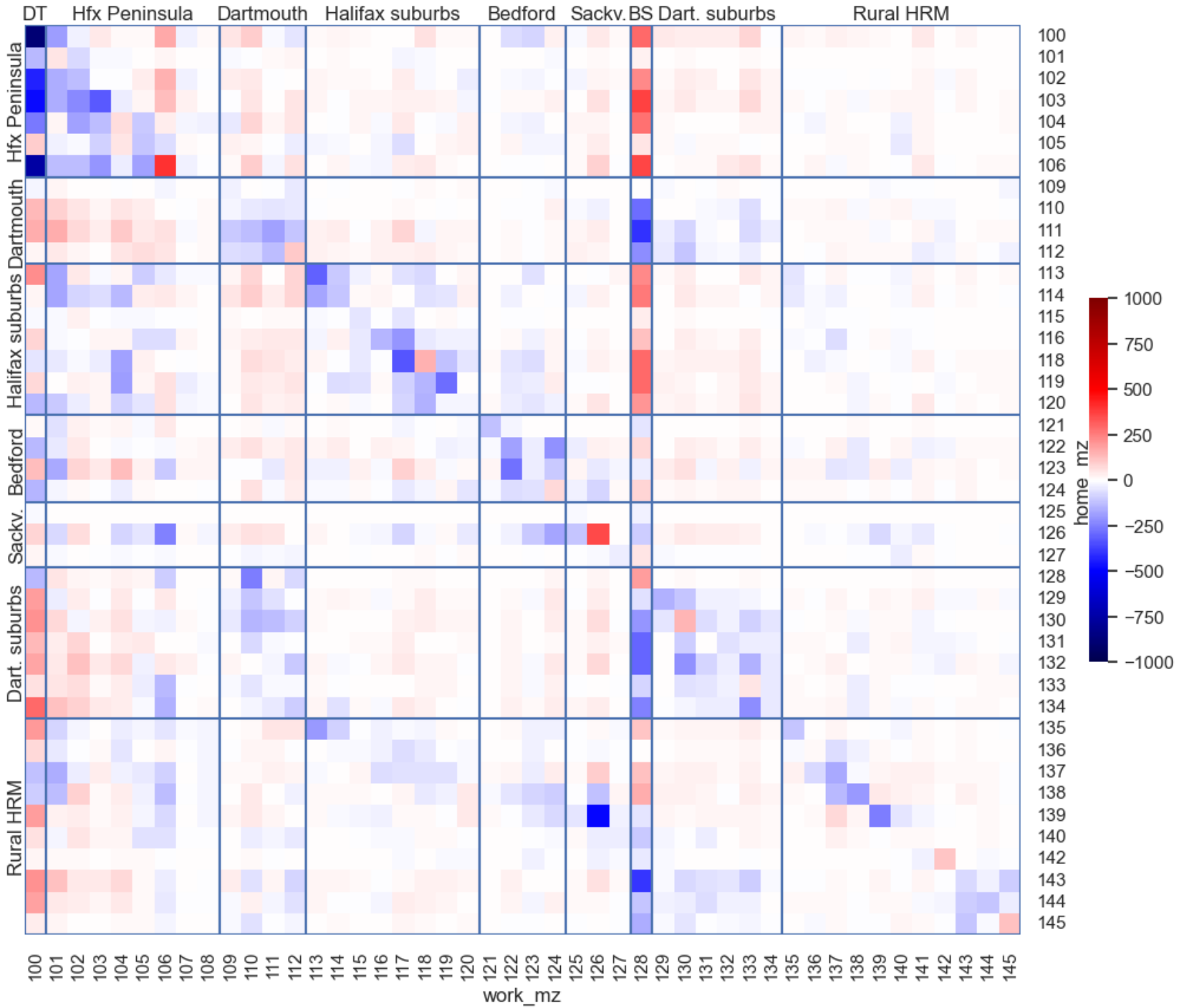


PCA K-factor matrix (clipped to [1/2, 2])



Difference in predicted flows, model vs. observed

PCA K-factor scaled model



Difference in predicted flows, model vs. observed

Uncalibrated model (Toronto parameters)

[Signed] GEH statistic for difference in predicted flows, model vs. observed

PCA K-factor scaled model

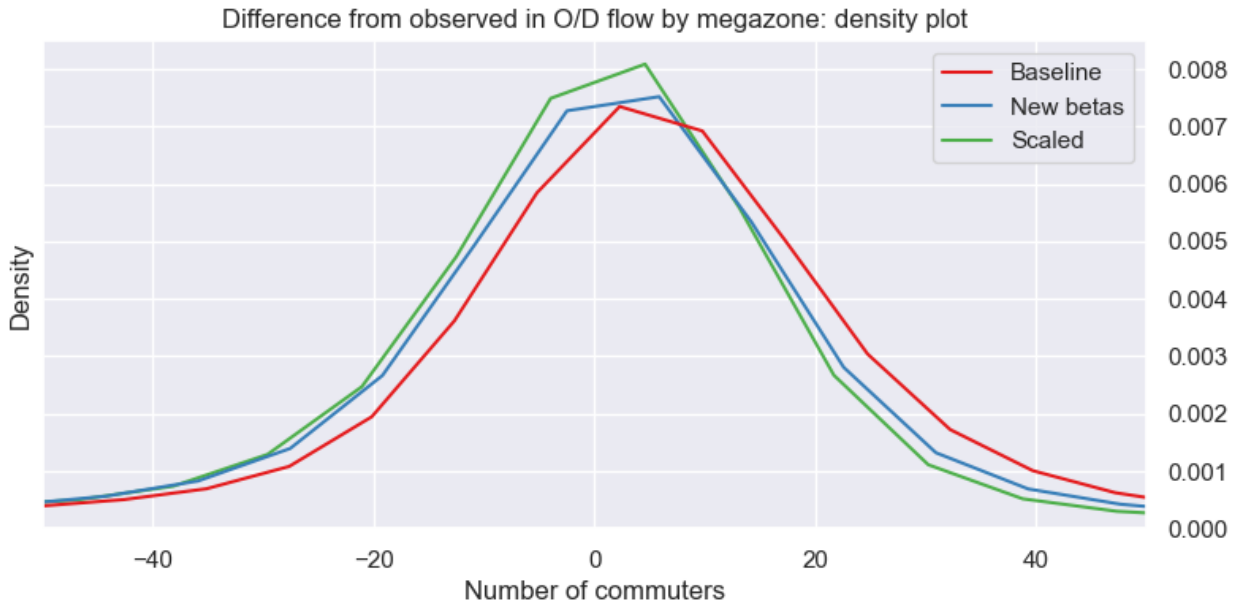
$$GEH = \sqrt{\frac{2(M-C)^2}{M+C}}$$

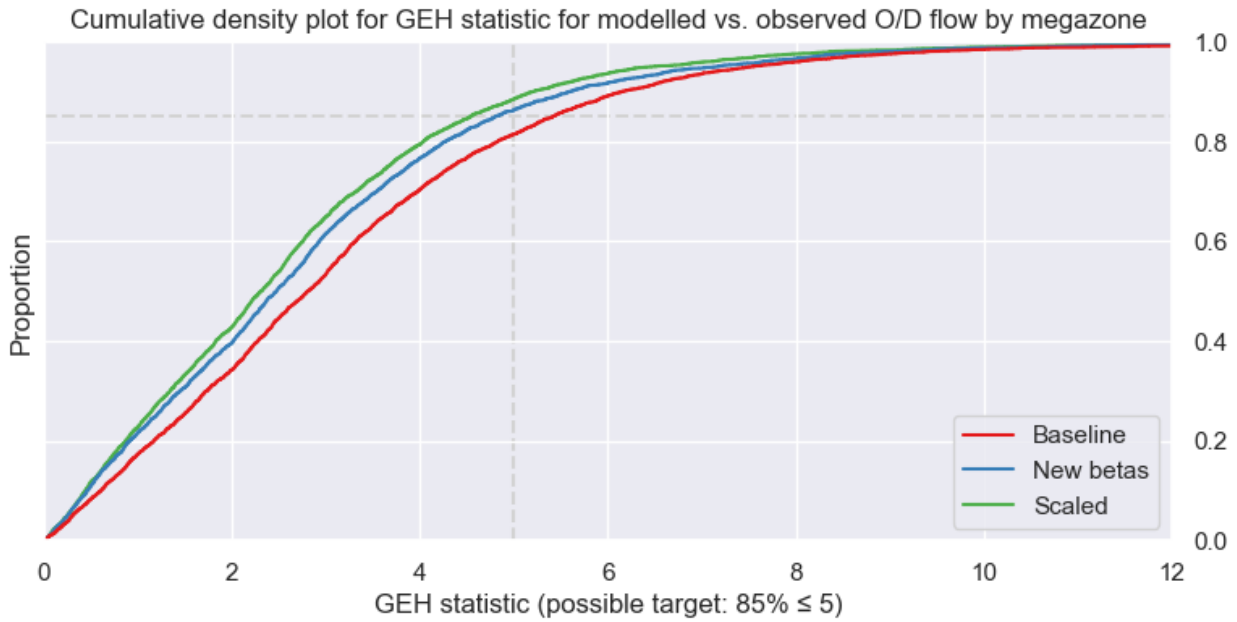
[Signed] GEH statistic for difference in predicted flows, model vs. observed

Uncalibrated model (Toronto parameters)

$$GEH = \sqrt{\frac{2(M-C)^2}{M+C}}$$

Does this improve the fit?





Questions? Suggestions?